Technical Report of the Survey of Adult Skills (PIAAC)

OECD

2013

Table of Contents

Foreword: The Programme for International Assessment of Adult Competencies – An Overview

Irwin Kirsch, ETS; William Thorn, OECD

Policymakers have become increasingly concerned not only over the levels of traditional literacy skills in their populations but also the growing importance of human capital and the broadening of the skills that will be needed to sustain productivity and social cohesion. The increased importance of human capital and the learning that is associated with it has led to a critical need for information about the distribution of knowledge, skills and characteristics that are needed for full participation in modern societies. The Organisation for Economic Co-operation and Development (OECD), in recognition of this need, initiated the development and implementation of a new international comparative survey of adults named the Survey of Adult Skills, as part of its Programme for the International Assessment of Adult Competencies (PIAAC), with the following goals and objectives:

- provide policymakers in each participating country with a baseline profile of adults in their country in terms of the knowledge, skills and competencies that are thought to underlie both personal and societal success;
- assess the impact of these competencies on a variety of social and economic outcomes at the individual and aggregate levels;
- gauge the performance of education and training systems in generating the required competencies; and
- help clarify some of the policy levers that could contribute to enhancing competencies.

The recently released OECD Skills Strategy (OECD, 2012a) identifies three key areas for action by governments in developing policies on skills designed to support sustainable long-term growth and employment creation and contribute to a fairer distribution of income and opportunities.

- Developing relevant skills: Ensuring that the supply of skills is sufficient in both quantity and quality to meet current and emerging needs is a central goal of skills policies. Supply can be ensured by developing the right mix of skills through education and training and by influencing the flow of skills through attracting and retaining talent. Supply is not only responsive to demand; it can also have an important influence on demand.

- Activating skills: People may have skill but for a variety of reasons may decide not to offer them to the labor market. Individuals withdraw from the labor force for a range of reasons, including personal preferences, life circumstances, or the lack of financial incentives to work. Encouraging inactive individuals to enter or reenter the labor force can increase the skills base of an economy. This requires identifying inactive individuals, possibly retraining them, ensuring that the benefit system offers them financial incentives to enter or return to the labor market, and removing demand-side barriers to hiring.

- Putting skills to effective use: Investment in skills development by individuals and governments needs to be accompanied by policies that ensure that these skills are used effectively. Moreover, the match between the skills demanded in a job and those of the person doing the job has an impact on further skills development: unused skills tend to atrophy, while new skills are, to a large extent, developed informally, often through work experience.

The Survey of Adult Skills responds directly to these themes and represents one of the key sources of empirical evidence which is available to help understand these issues. In particular, PIAAC considerably enhances knowledge about the stock of skills in the population by providing direct measures of key skills in addition to traditional measures such as educational attainment and labor force experience. It also offers a rich tool for better understanding the processes through which skills are gained, lost and retained and the extent to which skills are effectively used to create value for the economy and individuals.

Features of PIAAC

PIAAC has been planned as an ongoing program of assessment. The first cycle of the assessment has involved two "rounds." The first (covered by this report) took place over the period of January 2008-October 2013. A second round involving nine additional countries began at the start of 2012 and will extend to May 2016.[1] The second cycle of the assessment is expected to take place over 2018-2023.

The main features of the first cycle of PIAAC are described below.

Skills assessed

PIAAC assesses three domains of cognitive skill:

- Literacy (including reading components)

- Numeracy

- Problem solving in technology-rich environments (PSTRE)

The assessments of literacy and numeracy were undertaken by all participating countries. The assessments of reading components and problem solving were optional elements of the

[1] The following countries are participating in PIAAC Round 2: Chile, Greece, Indonesia, Israel, Lithuania, New Zealand, Singapore, Slovenia and Turkey.

assessment in Round 1 of the study.[2] Of the countries that reported results in Round 1, most implemented the reading components assessment, with the exceptions being Finland, France and Japan. And most implemented problem solving, with the exceptions being France, Italy and Spain.

A brief overview of the domains of competence assessed in PIAAC is provided below. The conceptualization of these domains is explained in more detail in Chapter 2 (see also OECD, 2012b).

Literacy

Literacy is defined in PIAAC as: *"understanding, evaluating, using and engaging with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential"* (OECD, 2012b). "Literacy" in PIAAC does not include the ability to write or produce text, skills commonly falling within the definition of literacy.[3] However, at the same time, "literacy" is a broader construct than "reading," narrowly understood as a set of strategies for decoding written text. It is intended to encompass the range of cognitive strategies (including decoding) that adults must bring into play to respond appropriately to a variety of texts of different formats and types in the range of situations or contexts in which they read. A unique feature of the assessment of literacy in PIAAC is that it assessed adults' ability to read digital texts (e.g., texts containing hypertext and navigation features such as scrolling or clicking on links) as well as traditional print-based texts.

To provide more detailed information about adults with poor literacy, the assessment of literacy in PIAAC was complemented by a test of "reading component" skills. Reading components represent the basic set of decoding skills which provide necessary preconditions for gaining meaning from written text – knowledge of vocabulary, ability to process meaning at the level of the sentence, and fluency in the reading of passages of text.

Numeracy

Numeracy is defined in PIAAC as *"the ability to access, use, interpret and communicate mathematical information and ideas, in order to engage in and manage the mathematical demands of a range of situations in adult life"* (OECD, 2012b). Numeracy is further specified through the definition of "numerate behavior," which involves managing a situation or solving a problem in a real context by responding to mathematical information and content represented in multiple ways.

It is recognized that literacy skills such as reading and writing constitute an enabling factor for numerate behavior and that when mathematical representations involve text, performance on numeracy tasks is, in part, dependent on the ability to read and understand text. However, numeracy in PIAAC involves more than applying arithmetical skills to information embedded in text. In particular, numeracy relates to a wide range of skills and knowledge (not just arithmetic

[2] In Round 2, there were no optional components, so the assessments of reading components and problem solving were treated as core components.
[3] The practical difficulties of assessing writing skills in the context of an international assessment made it impossible to include this as part of the assessment.

knowledge and computation), a range of responses (which may involve more than numbers), and responses to a range of representations (not just numbers in texts).

Problem solving

In PIAAC, problem solving in technology-rich environments is defined as "*using digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks.*" The first wave of PIAAC focused on "*the abilities to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, and accessing and making use of information through computers and computer networks*" (OECD, 2012b).

The PSTRE domain of PIAAC covers the specific class of problems people deal with when using information and computer technology (ICT). These problems share the following characteristics:

- The existence of the problem is primarily a consequence of the availability of new technologies.
- The solution to the problem requires the use of computer-based artifacts (applications, representational formats, computational procedures).
- The problems are related to the handling and maintenance of technology-rich environments themselves (e.g., how to operate a computer, how to fix a settings problem, how to use the Internet browser in a technical sense).

PSTRE represents a domain of competence which involves the intersection of the set of skills that are sometimes described as "computer literacy" (i.e., the capacity to use ICT tools and applications) and the cognitive skills required to solve problems. Some knowledge of how to use basic ICT input devices (e.g., use of a keyboard and mouse and screen displays), file management tools, applications (word processing, email) and graphic interfaces is essential in order to be able undertake assessment tasks. However, the objective is not to test the use of ICT tools and applications in isolation, but rather to assess the capacity of adults to use these tools to access, process, evaluate and analyze information effectively.

Other information on skills

Literacy, numeracy and PSTRE constitute a subset of the skills and competencies that are demanded in the labor market and mediate access to resources and services more generally in society. Along with specific technical and professional skills, other generic skills such as communication, interaction (such as the capacity to relate to others and work cooperatively), skills related to learning and the transmission of knowledge, as well as physical skills are valued to a greater or lesser extent on the labor market. In order to provide a more complete picture of the skills endowment of the adult population, PIAAC collected a considerable amount of information on the skills possessed and used by adults in addition to the measures of proficiency in literacy, numeracy and PSTRE. This information was collected in the form of self-reports as these skills are, for the most part, difficult, if not impossible, to assess directly in an international comparative context or through population surveys.[4]

Qualifications and work experience

Educational qualifications and work experience are commonly used proxies for individuals' skill endowments. PIAAC collected information on respondents' highest level of educational attainment as well as regarding the duration of work experience and mobility. This was complemented with information on respondents' perceptions regarding the educational qualifications and work experience they believed are normally necessary to get the job they currently occupied as well as the qualifications needed to perform this job satisfactorily.

Use of skills at work

Information was collected from respondents regarding four broad categories of generic work skills: cognitive, interaction and social, physical and learning.[5] Cognitive skills encompass reading, writing, mathematics and the use of ICT. Interaction and social skills cover collaboration and cooperation, planning the work and time of one's self and others, communication and negotiation, and customer contact (e.g., selling products and services and advising). Physical skills involve the use of gross and fine motor skills. Learning skills cover activities such as the instruction of others, learning (formally or informally) and keeping up to date with developments in one's field of professional activity.

The approach used in PIAAC owes much to the Job Requirements Approach (JRA) pioneered in the UK Skills Survey (Felstead et al., 2007). The JRA method consists of asking individuals about the importance of different types of tasks performed at work and subsequently inferring the types of skills that are required from their answers. By focusing on job tasks, this approach is considered to provide a more objective description of these skills than an approach relying on subjective self-assessments by individuals of the type and level of skills they possess.

Respondents were also asked about the extent that they believe their skills (considered globally) match the requirements of the job in which they were currently working.

Work-related training

Given the importance of work-related training as a potential source of skills and as an element of a strategy for the maintenance and upgrading of workforce skills, information was collected on participation by respondents in training of both a formal and informal nature over the 12 months prior to the interview.

Personal characteristics, background and outcomes

The PIAAC background questionnaire (BQ) included a range of information regarding the factors which influence the development and maintenance of skills such as education, social background, engagement with literacy and numeracy and ICT (both in and outside of work),

[4] A framework for the measurement of teamwork was developed for the Adult Literacy and Lifeskills study, but was not considered robust enough for inclusion in an international comparative assessment (Murray, Clermont and Binkley, 2005). See Baethge and Arends (2009) for the results of a feasibility study of measures of vocational skill in an international comparative context.

[5] The exact questions can be found in OECD (n.d.)

language background. Information was also collected on outcomes which may be related to skills. This included the current activity of respondents, employment status and income. In terms of noneconomic outcomes, PIAAC included questions on health status, volunteering, political efficacy and social trust.

Test delivery

PIAAC was designed as a computer-based assessment (CBA) and was delivered on a laptop computer. The BQ was administered in a computer-assisted personal interview (CAPI) format by the interviewer. The cognitive assessment was taken by most respondents in the CBA format under the supervision of the interviewer. Respondents with no (or extremely limited experience) with the use of computers were given a pencil-and-paper version of the literacy and numeracy components of the assessment. Respondents with computer skills but who possessed poor literacy and numeracy skills were directed to the reading components test, which was taken in pencil-and-paper format only. However, interviewers timed the completion of the reading components tasks using the computer application.

Respondents took the assessment in their own homes or in another location to which the interviewer agreed. They were free to take as much or as little time as required to complete the test. However, interviewers were trained to encourage respondents that took an excessive amount of time to undertake the assessment or were obviously experiencing difficulties to move through the test or terminate it.

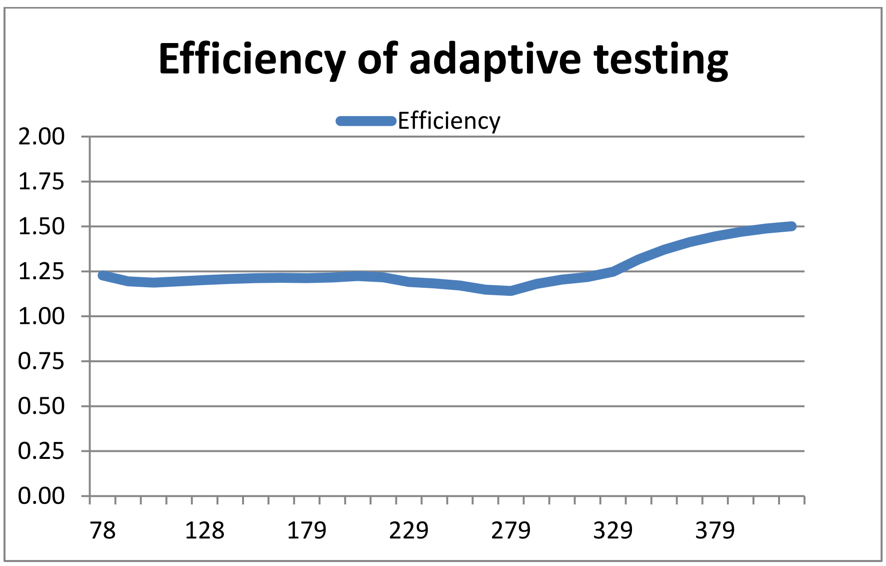
Adaptive testing

One outcome of introducing CBA in PIAAC was the use of adaptive algorithms to optimize the delivery of test items within a domain to estimated proficiency levels of individuals, thereby allowing PIAAC to provide more reliable information about skills in a relatively short period of time. Adaptive tests can be roughly distinguished as belonging to one of two groups: item-level adaptive tests and multistage adaptive tests. Item-level adaptive tests have been traditionally referred to as “computer adaptive tests” (CATs) and have been in vogue for some time. The idea of a CAT is intriguing and much research has been conducted; however, significant challenges remain. Perhaps the most important one is that CATs assume (in practically all cases) that multiple-choice items, or at best automatically scoreable short constructed response items, are used. Items that cannot be automatically scored are not usable in a CAT.

The multistage adaptive design used in PIAAC is a natural generalization of a CAT. It is an extension in the sense that the CAT algorithm “decides” on the choice of the next item after each response, whereas multistage algorithms allow the choice of the next cluster of items either after one or multiple responses. This provided more information and therefore the opportunity to accumulate greater accuracy in the decision. An additional advantage of a multistage CAT is that item types can be mixed – a multistage test can be designed to decide about the next cluster of items to be administered solely based on the automatically scored responses after a cluster of mixed item types has been administered. Moreover, using item clusters instead of individual items for adaptive decisions reduced the likely dependence of the stage adaptive selection on item-by-country interactions compared to the effects to be expected when using item-level adaptive tests.

Figure 1 shows the efficiency of the PIAAC literacy scale multistage adaptive test over a more traditional linear test using the same identical literacy item set defined as the ratio of two test information curves. The ratio of the two test information curves is shown on the vertical axis whereas the literacy scale is shown on the horizontal axis. As shown here, the adaptive test is 15 to 47 percent more efficient, which means that we can obtain the same amount of test information as we might expect from a test that is 15 to 47 percent longer. In addition, it should be noted that there is no proficiency range where adaptive testing is less informative. The success of using a multistage adaptive test design in PIAAC was largely due to being able to optimize the design, as we did not have any open-ended items that required human scoring and we had empirical evidence that the item parameters for trend items were identical regardless of the position of items in the assessment. This is not always the case with school-based comparative surveys.

Figure 1: Efficiency of the multistage adaptive testing model of the literacy scale used in PIAAC



Countries participating in PIAAC

In total, 28 countries participated in the first round of PIAAC at some point over 2008-2013, with 24 completing the Main Study and reporting results. The countries starting the study are listed in Table 1 together with whether they completed key phases of the study and reported results. Both Chile and Portugal only completed the Field Test.

Table 1: Participation in PIAAC (Round 1)

Country	Field Test completed	Main Study completed	Results reported
Australia	yes	yes	yes
Austria	yes	yes	yes
Canada	yes	yes	yes
Chile	yes	no	no
Cyprus [6]	yes	yes	yes
Czech Republic	yes	yes	yes
Denmark	yes	yes	yes
England/N. Ireland (UK)	yes	yes	yes
Estonia	yes	yes	yes
Finland	yes	yes	yes
Flanders (Belgium)	yes	yes	yes
France	yes	yes	yes
Germany	yes	yes	yes
Ireland	yes	yes	yes
Italy	yes	yes	yes
Japan	yes	yes	yes
Korea	yes	yes	yes
Netherlands	yes	yes	yes
Norway	yes	yes	yes
Poland	yes	yes	yes
Portugal	yes	no	no
Russian Federation [7]	yes	yes	yes
Slovak Republic	yes	yes	yes
Spain	yes	yes	yes
Sweden	yes	yes	yes
United States	yes	yes	yes

Chile, New Zealand, and Slovenia will continue to implement PIAAC as participants in the second round of the assessment.

In two countries participating in Round 1, PIAAC did not provide full national coverage of the adult population. In Flanders (Belgium), PIAAC was implemented only in the region of Flanders. In the UK, the assessment was undertaken in England and Northern Ireland only.

[6] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[7] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

The development and implementation of PIAAC

The process of the development and implementation of PIAAC can be seen as involving four broad phases: scoping, development, implementation, and data preparation and analysis.

The scoping phase (2002-2007)

Work within the OECD on a data development strategy regarding adult skills began in 2002 with the convening of an expert group on adult skills. A paper based on the conclusions of that meeting was presented to the OECD's Education and Employment, Labour, and Social Affairs committees in late 2003. The paper provided a rationale for an OECD strategy for the assessment of adult skills and identified four key issues for decision in the course of developing such a strategy:

- whether the strategy should be based on undertaking an assessment of the whole adult population or on a sequence of assessments targeted at different age groups,
- which competencies should be assessed,
- what relationship a program of adult assessment should have with the Programme for International Student Assessment (PISA), and
- what weight should be placed on trend data.

While not presenting any conclusions, the paper argued strongly that the implementation of a series of assessments targeted at particular population subgroups rather than an "omnibus" survey of the adult population should be considered. It also argued that the model of competence developed by DeSeCo (Rychen and Salganik, 2003) should guide selection of the domains of competence to be assessed.

In line with the recommendations of the paper, an international expert group (IEG) was established to develop an operational strategy for an international assessment of adult competencies over the following 18 months.

In October 2005, the IEG considered a strategy for PIAAC based on its work as well as on policy priorities identified by the OECD's education and employment policy committees. The main elements of this strategy were as follows:

- PIAAC was to constitute a multicycle program of assessment, with each cycle lasting five years. The first cycle of data collection would be scheduled for 2009 (or early 2010, depending on progress with the research agenda). PIAAC would survey a representative sample of the adult population between 16 and 65 years of age, including the nonemployed, in a household context and would provide the option of oversampling a cohort of young adults and/or older workers, and of resurveying the selected oversampled cohort(s) in subsequent cycles.
- The direct assessment would focus on the measurement of ICT-related competences, defined for the purpose of PIAAC as the capacity of individuals to access, manage, integrate, evaluate and reflect on information using modern technologies. This would be

accompanied by a short assessment of document literacy and an assessment of reading components to be taken by respondents with poor levels of literacy.

- Subsequent waves of the assessment would repeat administration of some components of the first to allow the establishment of trends. The development and implementation of new domains (e.g., an employer survey in 2014 and an assessment of interpersonal skills in 2019) would be a feature of the program.

The IEG broadly welcomed the proposed strategy but expressed the view there should be a balance between the assessment of ICT competencies and reading and numeracy in order to ensure the relevance of the assessment to all adults in OECD countries. It also underlined the need to ensure that the assessment would provide reliable information regarding the entire spectrum of proficiency of adults in OECD countries.

An amended strategy was subsequently presented to the OECD's education and labor committees. While the basic features of the original strategy remained, in the revised version, the direct assessment component was conceived as an assessment of "literacy for the information age" rather than of ICT competencies. The balance of data collection was also shifted somewhat from the assessment of competencies towards the collection of information on other social and economic outcomes as well as contextual data that could be used to examine the development, functioning and impact of competencies.

In 2006, a series of expert papers were commissioned by the OECD covering topics relevant to the design of PIAAC. These included papers on planning for the direct assessment, the measurement of work-related training, adult learning, the description and discussion of approaches to the identification of the skill content of jobs using self-reports, school-to-work transition, and human capital and economic development. This work led, in particular, to the establishment of the basic features of the direct assessment in the form that would be subsequently implemented.

In particular, the concept of a single measure of "literacy for the information age" encompassing elements of reading, numeracy and problem solving as proposed in the 2005 strategy was replaced by the measure of three distinct domains – literacy, numeracy and PSTRE. The reporting of these domains as separate scales was proposed with the aim of facilitating interpretation of the results as well as facilitating linking PIAAC to the International Adult Literacy Survey (IALS) and the Adult Literacy and Lifeskills (ALL) survey.

Work began on the development of the proposed JRA module of PIAAC in 2007 and continued into early 2009. The objective was to develop and test around 15 minutes of questions relating to the task content of the main job held by the respondent (if employed) covering a range of the generic skills that were required in performing that job. Five countries agreed to participate in a pilot of the JRA: Australia, France, Greece, Korea and the United States.

The first draft in English of the pilot questionnaire and technical specifications for implementing the pilot were sent to participating countries at the end of May 2007. An extensive pretesting stage was then carried out. This involved carrying out cognitive interviews in each of the five participating countries to check on the wording of questions and the scales being used.

Piloting of the JRA module took place during 2008 and involved administration of the pilot questionnaire to a random sample of 500 employed persons as well as a sample of 100 primary-school teachers. The pilot questionnaire contained both JRA questions and a limited number of background questions on demographic and labor-market characteristics of respondents included to help establish the international comparability of the results. A series of country reports (written by national experts) plus a summary validation report (written by a consultant) was produced in the second half of 2008. The results were presented at an international validation seminar in early 2009 hosted by the European Centre for the Development of Vocational Training, or Cedefop.

Following a meeting of countries interested in participating in PIAAC in November 2007, a call for tender for services relating to the development and implementation of the first wave of PIAAC was finalized and released in late 2007 with a closing date of January 2008. Bids were sought for three distinct groups of services – the development of assessment instruments (Module 1), the development of the BQ and JRA (Module 2) and survey operations and project management (Module 3). A Consortium led by Educational Testing Service (ETS) of Princeton, NJ, involving institutions from the US, the Netherlands, Flanders (Belgium), Germany, and Luxembourg, was selected by the PIAAC Board of Participating Countries (BPC) to undertake all three modules.

Development phase (2008-2009)

The first phase of the implementation of PIAAC involved work in three main areas:

- development of the PIAAC assessment frameworks, the instruments and questionnaires, the delivery platform, and other IT tools and technical standards
- preparation of national versions of the instrumentation
- preparation for the Field Test

The development of frameworks for the new assessment domains in PIAAC (PSTRE and literacy components) and the updating of the frameworks for literacy and numeracy used in ALL for use in PIAAC largely took place during 2008. This work was guided by three subject matter expert groups – covering the domains of literacy, numeracy and PSTRE, respectively. Draft framework documents were reviewed by the BPC in October 2008 and the final versions approved in April 2009. The selection of items from IALS and ALL to serve as linking items in literacy and numeracy and the development of new items took place in parallel with the development of the frameworks. Final selection of items for the Field Test took place in March 2009.

Development of the BQ took place over 2008 and 2009, with the Field Test version being finalized in 2009. This was guided by the BQ Expert Group and also involved input from the other subject matter expert groups, particularly in relation to questions regarding the use of and engagement with literacy, numeracy and ICT. The BPC was also closely involved in the development process, reviewing the contents of the proposed BQ twice before its finalization in early 2009.

The PIAAC Technical Standards and Guidelines (TSG), which define the quality standards that were to be met throughout the process of the development and implementation of the assessment, were prepared over 2008 and early 2009. A first draft of the TSG was reviewed by the BPC in November 2008 and subsequently by the Technical Advisory Group (TAG). A final version (which incorporated comments made by the BPC and the TAG) was agreed upon by the BPC in April 2010. A final version of the TSG was released in December 2010 for the Field Test and revised in December 2012 for the Main Study.

A major challenge in developing PIAAC was building a test delivery application for use on a laptop computer that combined a CAPI application for administering the BQ and a CBA application for administering the direct assessment that could be released in over 30 different country and/or language versions. Initial versions of the CAPI application, the Virtual Machine (VM) and the cognitive modules were released in 2009. National versions of the delivery platform (in national test languages) for use in the Field Test were released for testing by countries in February-March 2010. Countries tested the platform using predefined scenarios. Two rounds of testing were undertaken. Reported problems were evaluated in terms of their potential impact on quality of the data from the Field Test and either fixed in subsequent releases of the VM prior to the Field Test or identified as a problem to be fixed in the Main Study version of the VM.

Participating countries were responsible for the translation and adaptation of the master English language versions of the BQ and cognitive instruments into the national survey languages. Translations were undertaken using a specially developed tool to facilitate the loading of translations into the PIAAC delivery platform. Following review and verification, the approved national versions were loaded into the delivery platform to create national versions of the PIAAC VM – the application running the assessment.

Implementation (2010-2012)

The Field Test data collection took place from April-June 2010. Twenty-six countries participated in the Field Test. Analysis of the outcomes of the Field Test was undertaken from October to early December 2010. The conclusions of this analysis and the overall assessment of the quality of the data from the Field Test were presented along with recommendations regarding the items to be included in the Main Study BQ and instruments to the TAG, the subject matter expert groups, NPMs and the BPC in a series of meetings in December 2010. Following their approval by the BPC, the necessary changes to the BQ and cognitive instruments were implemented by countries and verified by the international Consortium.

Main Study versions of national VMs were released to countries for testing starting in March 2011. Two rounds of testing took place. Final Main Study VMs were released in May 2011.

The main data collection was scheduled to take place over the period August 2011-March 2012. Twenty-two countries took part in this phase of the study. Most countries completed data collection at the end of March 2012 as planned. A number of countries extended the data collection period by varying durations to improve response rates. Two countries collected data on different timetables. Canada started collection in November 2011 to avoid having PIAAC in the field at the same time as the Canadian census and completed collection in June 2012. France undertook the main data collection over the period September-December 2012.

Data preparation, analysis and reporting (2012-2015)

All but two of the participating countries submitted national datasets to the Consortium from the end of May to the end of August 2012. France and the Russian Federation[8] submitted their data in 2013. Cleaning, weighting and scaling were undertaken in the second half of 2012. Scaled national datasets were released to countries in January 2013 for review. Final datasets were released in April 2013 and loaded into a tool called the Data Explorer. From this point, participating countries had access to anonymized[9] output from the international dataset through the Data Explorer in addition to their own data to allow preparation of national reports on PIAAC.

Following the release of the national databases in January, the public-use dataset and associated documentation were produced for release in October 2013.

Planning for the analysis and reporting of the results from PIAAC began at the end of 2009 when the BPC discussed a first draft outline of the contents of the first international report. Further discussions regarding the contents of the report took place from 2010 to 2012, informed by presentations of some exploratory analyses of the data from the Field Test. A final outline was approved in May 2012.

The first international report was written from September 2012 to July 2013 by a team from the OECD Secretariat with the assistance and support of the Consortium. A first draft of the report was reviewed in May 2013 by participating countries and an external panel of reviewers. The final draft was reviewed by countries in June 2013.

Analysis of the data from PIAAC by the OECD will continue over 2014-2015 with the release of a series of reports addressing some of the issues of particular interest to countries participating in PIAAC.

Relationship to previous surveys

PIAAC is the third of a series of international adult skills surveys which have been implemented since the mid-1990s by OECD countries. It was preceded by IALS (1994-98) and ALL (2003-06).[10]

Table 2 presents the skill domains assessed in the three assessments. Shading indicates that the assessments in these domains can be linked across surveys.

[8] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[9] Countries were identified by codes rather than actual names.

[10] See OECD and Statistics Canada (2000), Statistics Canada and OECD (2005), and OECD, and Statistics Canada (2011) for information on the methods and results of IALS and ALL.

Table 2: Skills Assessed in PIAAC, ALL and IALS

PIAAC (2012)	ALL (2003-2006)	IALS (1994-1998)
Literacy (combined prose and document)	Literacy (combined prose and document*)	Literacy (combined prose and document*)
	Prose literacy	Prose literacy
	Document literacy	Document literacy
Reading components		
Numeracy	Numeracy	
		Quantitative literacy
Problem solving in technology-rich environments		
	Problem solving	

*Rescaled to form a single literacy scale combining the former separate prose and document literacy scales.

IALS assessed three domains of literacy – prose literacy, document literacy and quantitative literacy. Prose literacy was defined as the knowledge and skills needed to understand and use *continuous* texts – information organized in sentence and paragraph formats. Document literacy represented the knowledge and skills needed to process documents, or *information organized in matrix structures* (i.e., in rows and columns). The type of documents covered by this domain included tables, signs, indexes, lists, coupons, schedules, charts, graphs, maps and forms. Quantitative literacy covered the skills needed to undertake arithmetic operations such as addition, subtraction, multiplication or division either singly or in combination using numbers or quantities embedded in printed material.

The major change between IALS and ALL was the replacement of the assessment of quantitative literacy with that of numeracy and the introduction of the assessment of problem solving. Numeracy represented a broader domain than that of quantitative literacy, covering a wider range of quantitative skills and knowledge (not just computational operations) as well as a broader range of situations in which actors had to deal with mathematical information of different types (not just situations involving numbers embedded in printed materials) (Gal, et al., 2005, p.151). Problem solving was defined as "goal-directed thinking and action in situations for which no routine solution procedure is available" (Statistics Canada & OECD, 2005, p.16).

PIAAC has been designed to link to IALS and ALL in the domain of literacy and ALL in numeracy. To ensure strong links in literacy and numeracy with IALS and ALL, approximately 60% of the assessment items in these two domains in PIAAC have been drawn from these previous surveys.

In the domain of literacy, PIAAC differs from IALS and ALL in two main ways. First, literacy is assessed on a single scale rather than on two separate (prose and document literacy) scales. For the purposes of comparison, the results of IALS and ALL have been rescaled on the PIAAC literacy scale. Second, while the measurement framework for literacy in PIAAC draws heavily on those used in IALS and ALL, it expands the kinds of texts covered to include electronic and combined texts in addition to the continuous (prose) and noncontinuous (document) texts of the IALS and ALL frameworks. In addition, the assessment of literacy was extended to include a measure of reading component skills which was not included in previous assessments.

The domain of numeracy remains largely unchanged between ALL and PIAAC. PSTRE constitutes a new domain. While it has some relationship to problem solving as conceived in ALL, the emphasis is on the skills necessary to solve "information problems" and the solution of problems in an ICT context rather than on analytic problem skills per se.

Comparability between background questions

The PIAAC BQ differs in a number of areas from the background questionnaires of IALS and ALL. In particular, the PIAAC BQ seeks more information about the use of skills in the workplace than does either IALS or ALL. In key areas such as educational attainment and labor-force status, the information in PIAAC and IALS and ALL is sought using comparable questions.

Countries participating in PIAAC and previous adult surveys

In total, 17 of the countries participating in PIAAC (Round 1) participated in either IALS, ALL or both (see Table 3 below), with 16 countries participating in IALS, seven in ALL and six in both. Results for France from IALS and for Korea from ALL have never been reported.

Table 3: Countries in Round 1 of PIAAC – Participation in IALS and ALL

	IALS			**ALL**	
Country	**1994**	**1996**	**1998**	**2003**	**2006**
Australia		X			X
Austria					
Canada	X			X	
Cyprus[11]					
Czech Republic			X		
Denmark			X		
Estonia					

[11] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 3 (cont.): Countries in Round 1 of PIAAC – Participation in IALS and ALL

	IALS			ALL	
Country	**1994**	**1996**	**1998**	**2003**	**2006**
England (UK)		X			
Finland			X		
Flanders (Belgium)		X			
France	X*				
Germany	X				
Ireland					
Italy			X	X	
Japan					
Korea				X*	
Netherlands	X				X
Northern Ireland (UK)		X			
Norway			X	X	
Poland	X				
Russian Federation[12]					
Slovak Republic					
Spain					
Sweden	X				
United States	X			X	

* Results not reported

As can be seen from Table 3, IALS was undertaken in three separate waves with data collection occurring in 1994, 1996 and 1998, and ALL was undertaken in two waves with data collection taking place in 2003 and 2006. Table 4 shows the number of observations of the performance in literacy and numeracy available for countries which undertook IALS or ALL prior to PIAAC as well as the period between observations. This varies significantly between countries in the case of literacy, depending on whether a country participated in IALS only or both IALS and ALL.

Table 4: Participation in literacy and numeracy assessments, dates of and periods between observations

Country	Domain	Observations	Date of survey	Years between observations
Australia	Literacy	3	1996, 2006, 2011	10, 5
Australia	Numeracy	2	2006, 2011	5
Canada	Literacy	3	1994, 2003, 2011	9, 8
Canada	Numeracy	2	2003, 2011	8
Czech Republic	Literacy	2	1998, 2011	13

[12] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Country	Domain	Observations	Date of survey	Years between observations
Denmark	Literacy	2	1998, 2011	13
England (UK)	Literacy	2	1996, 2011	15
Finland	Literacy	2	1998, 2011	13
Flanders (Belgium)	Literacy	2	1994, 2011	17
Germany	Literacy	2	1994, 2011	17
Italy	Literacy	3	1998, 2003, 2011	5, 8
Italy	Numeracy	2	2003, 2011	8
Netherlands	Literacy	3	1994, 2006, 2011	12, 5
Netherlands	Numeracy	2	2006, 2011	5
Northern Ireland (UK)	Literacy	2	1996, 2011	15
Norway	Literacy	3	1998, 2003, 2011	5, 8
Norway	Numeracy	2	2003, 2011	8
Poland	Literacy	2	1994, 2011	17
Sweden	Literacy	2	1994, 2011	17
United States	Literacy	3	1994, 2003, 2011	9, 8
United States	Numeracy	2	2003, 2011	8

Management structure

The development and implementation of PIAAC was steered by the BPC. The BPC is formally constituted as a body of the OECD and its role is defined by a mandate approved by the OECD Council. OECD countries participating in PIAAC are automatically members of the BPC. Non-member countries participating in PIAAC could be invited to join the BPC. With two exceptions, Cyprus[13] and the Russian Federation[14], all countries participating in the first round of PIAAC are members of the BPC. While countries have only one vote on the BPC, most are represented on the BPC by delegates from both ministries of labor and education.

The BPC is the main decision-making body regarding PIAAC with responsibility for setting priorities for the project, developing a program of work and budget, monitoring the implementation of the program of work, and evaluating its impact and disseminating results. It usually meets twice a year. All key elements of the design of PIAAC, its implementation and the reporting of results were reviewed and approved by the BPC. Decisions which needed to be made on a timetable that did not fit the BPC's meeting schedule were made through a process of written procedure.

The BPC reports to the Education Policy Committee (EDPC) and the Employment, Labour and Social Affairs Committee (ELSAC) of the OECD. It consults with these two bodies regarding policy priorities for PIAAC and reports to them on the progress of PIAAC on a regular basis.

[13] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[14] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

The budget and program of work of PIAAC (and any changes to it) were agreed upon by the two committees before submission to the OECD Council for approval.

The OECD Secretariat is responsible for supporting and advising the BPC and for ensuring that the work program of the BPC and its decisions are implemented. In particular, the OECD Secretariat managed the contract with the Consortium covering the development and international component of the implementation of PIAAC. It was also responsible for the preparation of the international comparative report.

The Consortium was headed by ETS, which reported directly to the OECD and had responsibility for each of the subcontractors, plus the TAG and the subject matter expert groups. Other contractors working on PIAAC included cApStAn, DIPF (the German Institute for International Educational Research), GESIS (Leibniz Institute for the Social Sciences), IEA-DPC (the International Association for the Evaluation of Educational Achievement-Data Processing Center), ROA (the Research Centre for Education and the Labour Market) and Westat. Each organization had particular areas of responsibility associated with the development of the instruments and delivery platform; the development of operational procedures and standards; translation verification quality assurance and quality control; the support of countries in key areas such as sampling, scoring, interview training and platform testing, undertaking data processing, scaling and data analysis; as well as the preparation of data analysis tools.

National implementation of PIAAC was managed by a range of organizations within participating countries. These included national statistical offices, public or private research and survey organizations contracted to manage implementation, government ministries, public research institutes and universities. In each participating country, the team responsible for the implementation of PIAAC was headed by a National Project Manager (NPM). Participating countries were responsible for aspects of survey implementation such translation and adaptation, sampling, data collection, scoring and coding and preparation of their national data base.

Close contact was maintained between the Consortium and national implementation teams throughout the project. Meetings of NPMs were held on a regular basis over the life of the project (approximately two meetings per year) and were attended by all participating countries. These constituted forums for the provision and exchange of information, the delivery of training and discussion of progress with the project and matters of concern raised by countries. The Consortium was responsible for managing NPM meetings. The OECD Secretariat was present at meetings and provided a regular update on discussions and decisions at the BPC as well as other relevant issues.

Organization of the report

This technical report was written by members of the consortium and is organized into six sections.

Section One: This contains four chapters that focus on assessment design, development of the cognitive instruments, development of the BQ, and the adaptation, translation and verification of the complete set of survey materials.

Section Two: This includes five chapters, with three dealing with development of the functionality to support development of the cognitive items. It also has a chapter covering development of the CAPI questionnaire software including the authoring tool and data export formats. In addition, it has a chapter focusing on the development and testing of the integrated computer platform that was used to deliver both the Field Test and main survey instruments.

Section Three: This consists of four chapters that cover field operations, quality control, scoring reliability and data management. Field operations include issues dealing with staffing, field management, production and response rates, and contact and outreach. Quality control includes activities that were undertaken prior to, during and after data collection during both the Field Test and the Main Study. Scoring focuses on preparing countries to score their paper-and-pencil cognitive booklets as well as to code open-ended questions in the BQ. It also deals with the design and procedures associated with obtaining estimates of within and between country interrater agreements. The chapter on data management covers data management systems, manuals and training that were provided to countries, as well as the tasks and responsibilities of each national center as well as the responsibilities and tasks conducted by the Consortium.

Section Four: This contains three chapters which focus on topics associated with sample design, survey weighting and variance estimation and indicators of overall sample quality.

Section Five: This is the largest section in the report, containing seven chapters. Theses cover data analysis and the preparation of the data products. Included are chapters describing the approach taken to scaling the cognitive data, evaluating the scaling outcomes and creating the proficiency scales for the cognitive domains. Other chapters deal with the validation of the BQ the creation of derived variables that are used in the analyses and that are available through the data products. Others cover the process of working with the expert groups to create described proficiency levels, reporting the results, and the development and use of data analysis tools.

Section Six: A set of appendices is provided here to help in understanding and using the PIAAC data.

References

Baethge, M., & Arends, L. (2009). *Feasibility Study VET-LSA: A comparative analysis of occupational profiles and VET programmes in 8 European countries – international report*. Bonn, Germany: Federal Ministry of Education and Research,

Felstead, A., Gallie, D., Green, F., & Zhou, Y. (2007). *Skills at Work, 1986 to 2006*. Oxford: ESRC Centre on Skills, Knowledge and Organisational Performance.

Gal, I., van Groenestijn, M., Manly, M., Schmitt, M. J., & Tout, D. (2005). Adult numeracy and its assessment in the ALL survey: A conceptual framework and pilot results. In S. Murray, Y. Clermont, and M. Binkley (Eds.), *Measuring adult literacy and life skills: New frameworks for assessment* (Catalogue 89-552-MIE, No. 13). Ottawa, Canada: Statistics Canada.

Murray, S., Clermont, Y., & Binkley, M. (Eds.) (2005). *Measuring adult literacy and life skills: New frameworks for assessment* (Catalogue No. 89-552-MIE, No. 13). Ottawa, Canada: Statistics Canada.

Organisation for Economic Co-operation and Development. (n.d.). *PIAAC background questionnaire*. Retrieved from http://www.oecd.org/edu/48442549.pdf

Organisation for Economic Co-operation and Development. (2012a). *Better skills. Better jobs. Better lives. A strategic approach to skills policies*. doi: 10.1787/9789264177338-en Organisation for Economic Co-operation and Development. (2012b), *Literacy, numeracy and problem solving in technology-rich environments: Framework for the OECD Survey of Adult Skills*. doi: 10.1787/9789264128859-en

Organisation for Economic Co-operation and Development and Statistics Canada. (2000). *Literacy in the information age: Final report of the International Adult Literacy Survey*. Paris and Ottawa, Canada: Authors.

Organisation for Economic Co-operation and Development and Statistics Canada (2011). *Literacy for life: Further results from the Adult Literacy and Life Skills Survey*. doi: 9789264091269-en

Rychen, D., & Salganik, L. (Eds.) (2003). *Key competencies for a successful life and a well-functioning society*. Göttingen, Germany: Hogrefe and Huber Publishers.

Statistics Canada and Organisation for Economic Co-operation and Development. (2005). *Learning a living: First results of the Adult Literacy and Life Skills Survey*. Ottawa, Canada, and Paris: Authors.

Note to Readers

General note

Throughout this report "PIAAC" refers to the Survey of Adult Skills (PIAAC). This differs from the terminology used in the *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills* and *The Survey of Adult Skills: Reader's Companion* in which the assessment undertaken over 2008-2013 is referred to as the Survey of Adult Skills (PIAAC) and "PIAAC" refers to the program of activities of which the survey is a product.

* * *

Cyprus

Readers should note the following information provided by Turkey and by the European Union Member States of the OECD and the European Union regarding the status of Cyprus:

A. Note by Turkey: The information in this document with reference to "Cyprus" relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the Island. Turkey recognizes the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Turkey shall preserve its position concerning the "Cyprus issue."

B. Note by all the European Union Member States of the OECD and the European Union: The Republic of Cyprus is recognized by all members of the United Nations with the exception of Turkey. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

* * *

Russian Federation

The data from the Russian Federation is *preliminary* and may be subject to change. Users should note that the sample for the Russian Federation does not include the population of the Moscow municipal region. The data published, therefore, do not represent the entire resident population aged 16-65 years in Russia but rather the population of Russia *excluding* the population residing in the Moscow municipal area. More detailed information regarding the data from the Russian Federation as well as that of other countries can be found in the *Technical Report of the Survey of Adult Skills*.

Acronyms

The following is a list of acronyms used throughout this report.

ALL	Adult Literacy and Lifeskills Survey
BPC	Board of Participating Countries
BQ	Background Questionnaire
CAPI	Computer-Assisted Personal Interview
CBA	Computer-Based Assessment
IALS	International Adult Literacy Survey
ICT	Information and Computer Technology
IRT	Item Response Theory
JRA	Job Requirements Approach
NPM	National Project Manager
OECD	Organisation for Economic Co-operation and Development
PBA	Paper-Based Assessment
PIAAC	Programme for the International Assessment of Adult Competencies

Chapter 1: PIAAC Assessment Design

Irwin Kirsch and Kentaro Yamamoto, ETS

The heart of any large-scale comparative survey is the assessment design. This chapter provides an overview of both the Field Test and Main Study designs. These designs were complex because PIAAC measured four domains – literacy, numeracy, reading components and problem solving in technology-rich environments – across two modes of administration – paper-and-pencil and computer delivered – while also offering participating countries both core and optional components. As the intent of PIAAC was to have its results linked to previous international adult assessments, these designs assumed that 60 percent of the literacy and numeracy tasks would come from ALL and IALS. New items were also developed for the literacy and numeracy domains and new measures developed for reading components and problem solving in technology-rich environments based on their respective frameworks.

The assessment designs assumed approximately 30-40 minutes of administration time for the BQ and JRA and 60 minutes for the direct assessment. The JRA items collected information on skill use at work, while the BQ collected contextual information about respondents, including their demographic characteristics, educational background, labor market experiences, and skill use outside of work. The JRA and background items were collected and processed through the use of a CAPI system. The target population ranged from 16 to 65 years of age.

1.1 Field Test goals and design

Field Tests are an integral part of any large-scale assessment and must be designed to yield adequate information relating to four key areas: survey operations, instrument quality, computer-delivery platform, and scaling and psychometric characteristics. Standardized procedures and quality mechanisms were embedded into various phases of PIAAC including survey development, implementation, and analysis and reporting of the data. The outcomes of the Field Test were used to assemble the final instruments for the Main Study and to modify or refine any of the operational issues detailed in the “standards and guidelines” document that improved the overall quality of the assessment.

1.1.1 Operational goals

Operation includes an examination of the efficiency and accuracy of data collection procedures, response rates for various subpopulations of interest, efficiency and accuracy of data processing including recoding, and data transmission. In particular, the following issues related to field operations needed to be examined:

- Review sample characteristics in terms of responses to BQ
- Review response rates by key background variables
- Evaluate coding of nonresponse interviews
- Identify and fix operational difficulties
- Summarize administration time for BQ as well as cognitive items
- Evaluate efficacy of scoring of paper-and-pencil items
- Evaluate efficacy of data capture
- Evaluate operational issues associated with International Standard Classification of Education coding and other BQ variables
- Evaluate efficacy and accuracy of data transmission
- Review and approve quality assurance mechanisms

1.1.2 Instrumentation

In addition to survey operations, the Field Test needed to provide quality information relating to the survey instruments, including adequacy of the scoring procedures, examination of translation and adaptation, and an evaluation of the scaling and analytic procedures that were used. In particular, the Field Test needed to address the following issues related to instrumentation:

- Review accuracy and comparability of survey instruments, including translation and scoring guides and all related manuals
- Evaluate the timing and flow of questions in the BQ
- Evaluate appropriateness of questions across participating countries
- Examine response distribution in all categories of BQ

1.1.3 Computer-delivery platform

PIAAC represents an innovation in large-scale assessment methodology in that the assessment was also computer based. PIAAC was the first large-scale assessment delivered on a laptop computer to respondents in their homes. An integrated computer-delivery platform was used to integrate the CAPI tool to be used for the administration of the BQ and the JRA with the tool that delivered the cognitive instruments. In its turn, the integrated PIAAC system needed to work in conjunction with the survey management systems of the organizations administering the survey in countries. Thus, in addition to looking at the instruments and survey operations, the Field Test also addressed the following issues related to the computer-delivery platform:

- Test and evaluate the functioning of the cognitive portion of the delivery platform, particularly response capturing and automatic scoring
- Test and evaluate the functioning of the CAPI system, particularly the flow of questions and efficiency of the system in capturing information
- Evaluate the accuracy of the interviewer's instructions
- Test the effectiveness of the system during the interview
- Verify the integration of the PIAAC platform with national survey management systems

1.1.4 Scaling and psychometric characteristics

The Field Test design allowed us to evaluate the psychometric characteristics of items and scales, including the evaluation of the equivalence of item parameters among linking items from IALS and ALL to PIAAC, and the equivalence of item parameters between paper-and-pencil and computer formats. In the case of PIAAC, the Field Test was also an opportunity to examine the role of computer familiarity and to determine the standards for branching respondents. In this regard, the Field Test provided initial IRT parameters that were used to construct the adaptive testing algorithm that were then implemented in the Main Study. In particular, the Field Test addressed the following issues associated with respect to IRT scaling and psychometric characteristics:

- Examine equivalence of item characteristics among the literacy and numeracy items common to IALS and ALL on the paper-and-pencil version
- Examine equivalence of item characteristics of literacy and numeracy items common to paper-and-pencil and computer-based formats
- Examine equivalence of item characteristics across languages within a country
- Examine equivalence of item characteristics across countries
- Identify tasks among the literacy, numeracy and problem-solving items that could be assembled into a core assessment
- Examine the expected proportions of subsamples routed to the different formats and to the different stages of the computer-delivered testlets based on preliminary background information and the core.

- Evaluate the overall psychometric characteristics and quality of the Field Test items to guide the selection of items for the Main Study

Because this was the first cycle of PIAAC, the Field Test was also viewed as a “dress rehearsal” of all newly developed aspects of the survey. In terms of sampling procedures, the Field Test did not need to be a full probability sample. However, critical aspects of sampling (such as sampling individuals within households), as well as other aspects of the overall sampling plan (such as descriptions of the sampling frames), and sampling guidelines had to be tested in this phase of the project. All quality control forms and procedures were also developed and tested. Finally, even though weights were not required for the Field Test, the weighting process was evaluated using the Field Test data.

1.2 The Field Test design: An integrated approach

This central Field Test design provided good item-level information on the full range of direct assessment measures included in PIAAC and was extremely useful in addressing other operational and psychometric issues identified above. The BQ and a core set of questions focusing on ICT helped to ensure that respondents who reported no familiarity with computers were routed to the paper-and-pencil version of the assessment. In order to link the paper-and-pencil and the computer-delivery formats, the remaining adults (the majority of adults in each country who are expected to pass the core) were randomly assigned to either the paper-and-pencil or computer-delivered branches of the Field Test (see Figure 1.1).

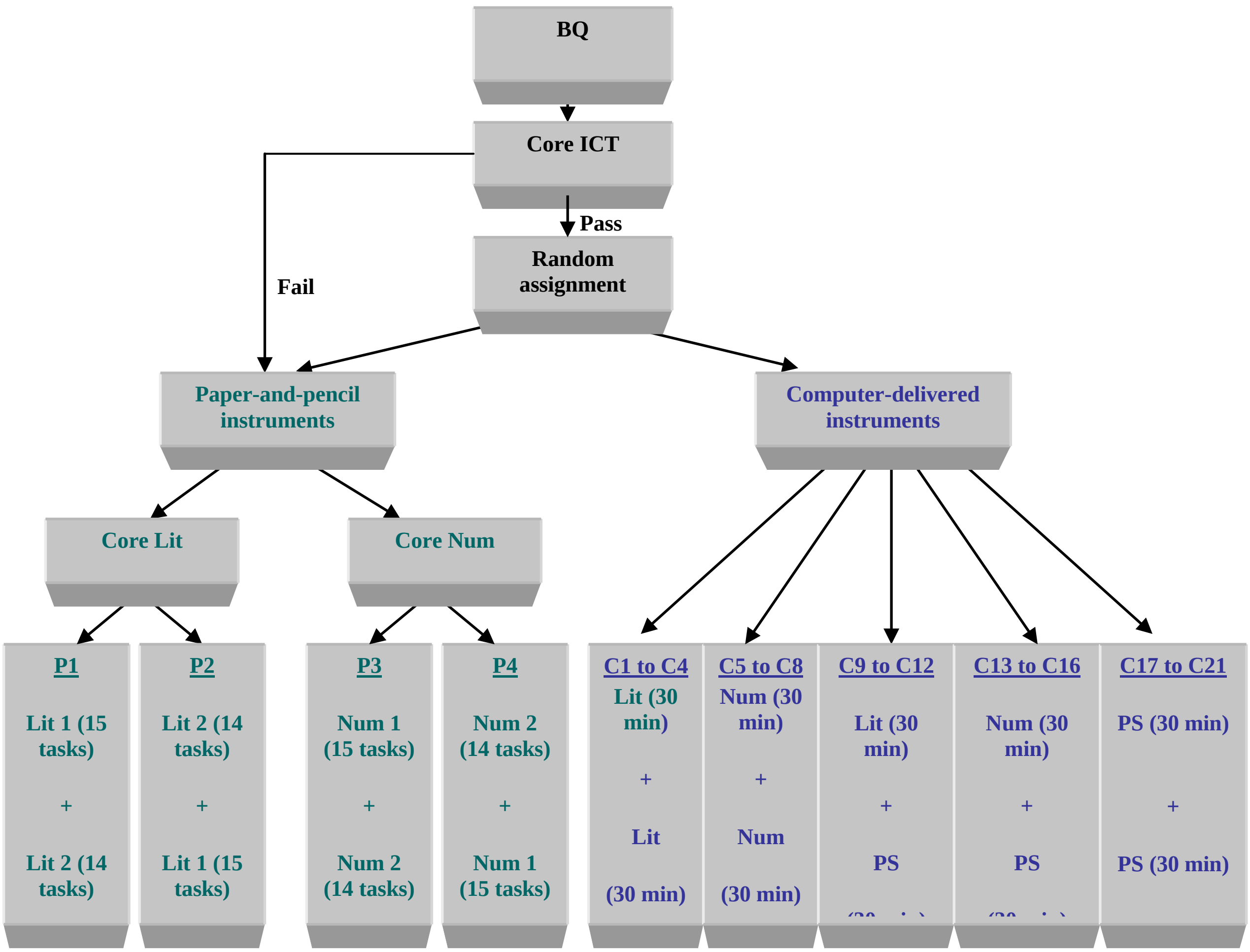
The BQ, including the JRA, was delivered by the interviewer using a computer-assisted format with respondents taking one of three variants, 30-40 minute variants (a 20-minute core set of items and one of three, 10-minute subsets) that were administered along with the cognitive instruments. The paper-and-pencil branch of the direct assessment was composed of a 10-minute core of either literacy or numeracy skills with six tasks each. This was followed by a pair of 20-minute clusters of literacy or numeracy, totaling 29 tasks, and a final 10-minute cluster of component skills. Four paper booklets were designed (details in Annex A1). Thus, each of the four direct assessment Field Test booklets was estimated to take 60 minutes.

In contrast to the paper-and-pencil branch of the Field Test design, the computer-delivered branch included 21 testlets that were 60 minutes long, consisting of a pair of 30-minutes blocks of items in each testlet[1] (as shown in Figure 1.4). As reflected in this design, each of the computer-delivered testlets contained only literacy tasks, only numeracy tasks, both literacy and problem-solving tasks, both numeracy and problem-solving tasks, or only problem-solving tasks. Overall, for the Field Test, there were 13 blocks that are 30 minutes long, grouped to form the 21 testlets: four blocks of literacy tasks (L1-L4), four blocks of numeracy tasks (N1-N4) and five blocks of problem solving tasks (PS1-PS5), as illustrated in Annex A2. The administration of these 21 testlets followed the administration of the BQ, including the ICT core as described above.

[1] The CBA comprised intact clusters of items that were grouped following a predetermined format. These groupings were not visible to users but are still called testlets for reference.

In this design, the direct assessment time was 60 minutes, on average, and each item was expected to be answered by a minimum of 150 adults based on an estimate of 1,500 respondents per country/per language (i.e., completed cases): 1,100 for the computer-delivered test and 400 for the paper-and-pencil test.

Figure 1.1: Paper-and-pencil Field Test assessment design, integrated



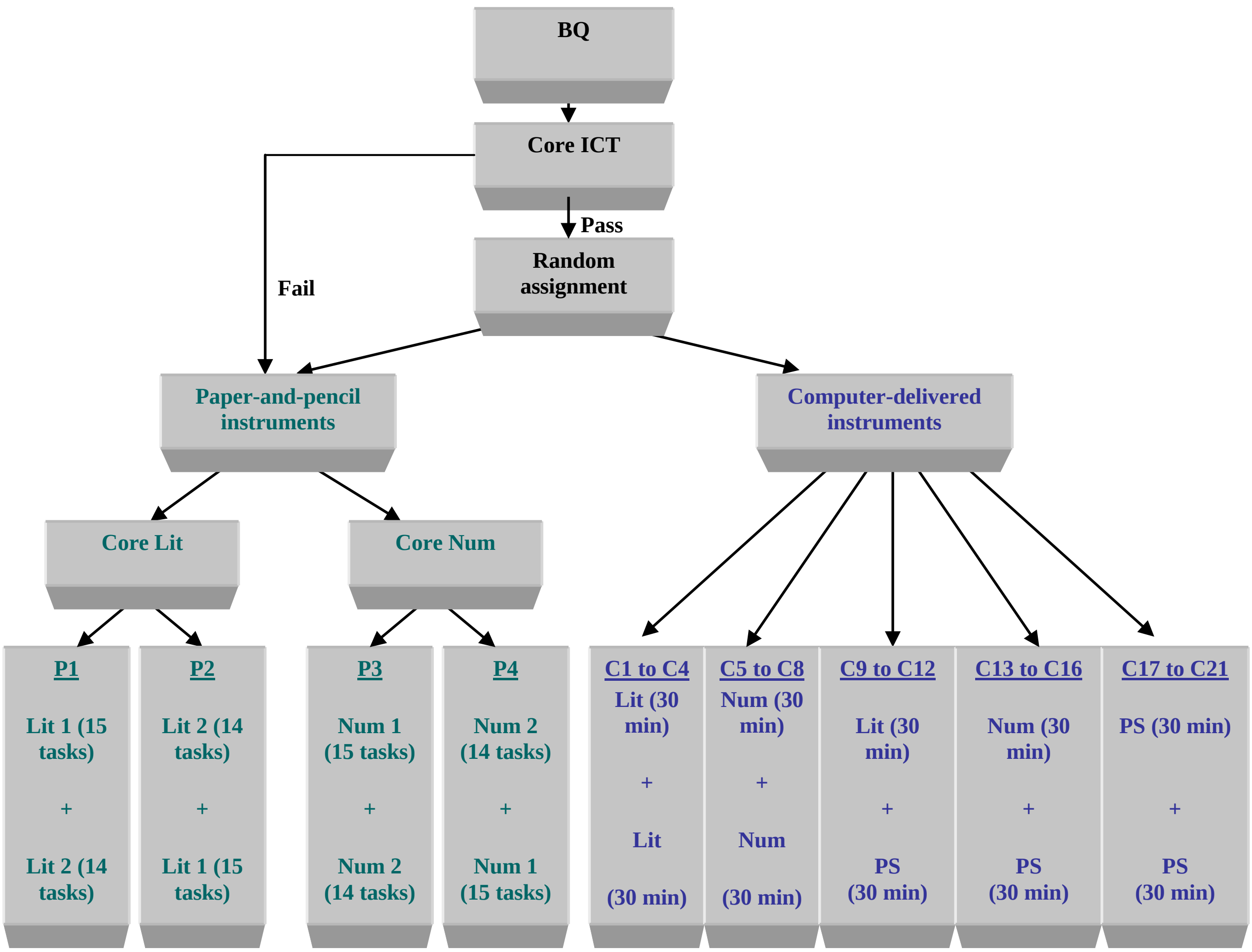
1.3 The role of international options in the Field Test assessment design

Some of the domains that were tested in the direct assessment were identified as international options. Which options were chosen by each of the participating countries had an impact on the Field Test and Main Study designs as well as on the required sample size.

1.3.1 Reading component skills as an international option

A country's decision not to assess reading components (one of the international options) had minimal impact on the overall Field Test design, as shown in Figure 1.2. Countries choosing not to include the reading components measures saved about 10 minutes in the overall assessment time and were able to reduce their sample size by a total of 100 adults. Under this design, assessment time was estimated to be 50 minutes, each item was expected to be answered by 150 adults, and the design was based on an estimate sample of 1,400 respondents per country/per language (i.e., completed cases): 1,100 who respond to the computer-delivered instruments and 300 who respond to the paper-and-pencil booklets.

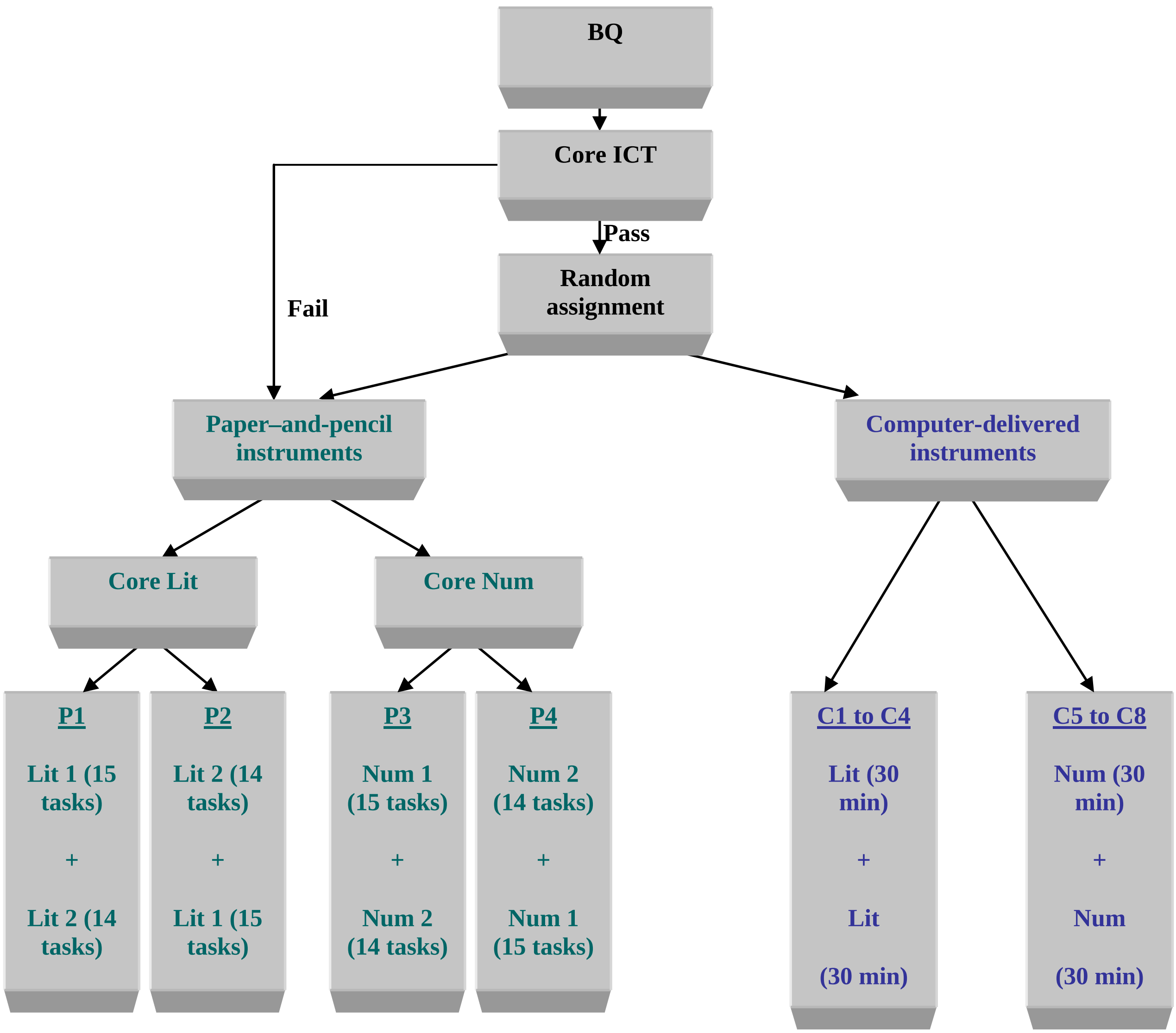
Figure 1.2: Paper-and-pencil Field Test assessment design, without reading components



1.3.2 Problem solving in technology-rich environments as an international option

The international option to include reading components but not to assess problem solving had a significant impact on both the sample size needed for the Field Test as well as on the number of computer-based booklets. This is shown in Figure 1.3. To compensate for the lack of covariance information, the number of respondents per item was increased but the overall sample size reduced by some 300 completed cases. In this design, assessment time per individual remained at 60 minutes, each item was answered by 200 adults, and was based on an estimate of 1,200 respondents per country/per language (i.e., completed cases): 800 who responded to the computer-delivered measures and 400 who responded to the paper-and-pencil items.

Figure 1.3: Paper-and-pencil Field Test assessment design, without problem solving



1.4 Item development needs

The item development requirements and goals for the literacy and numeracy domains are shown in Table 1.1. Overall, the Main Study required 24 items in each domain for the paper-and-pencil assessment and 48 items for the computer-delivered measures in each of the two domains. Of these, some 19 paper-and-pencil and 29 computer-delivered items were needed in each domain to serve as linking items. Linking items refer to items selected from IALS and ALL that were used to establish the link between PIAAC and these previous studies and between paper-and-pencil and computer-delivered formats. In order to meet these goals for each domain, it was necessary to develop and assess a larger pool of items for the Field Test.

The Field Test item pool required a total of 35 paper-and-pencil literacy and 35 paper-and-pencil numeracy items. The computer version needed 72 items for each domain. Of these, 42 were used to evaluate their utility as linking items for the computer-delivered measures while a subset of 25 was used to evaluate their utility for linking the paper-and-pencil and computer-delivered formats.

Table 1.1: Literacy and numeracy development item needs for PIAAC

Literacy or numeracy item development needs	Field Test		Main Study	
	Link	New	Link	New
Paper-based	25	10	19	5
Computer-based	42	30	29	19

As a new construct and domain for adult surveys, the assessment of problem solving in technology-rich environments involved scenarios of varying levels of complexities. Scenarios were designed to take between an average of five to 15 minutes to complete. Overall, 150 minutes of testing material was developed for the Field Test (approximately 16 scenarios of varying lengths) with some 75 minutes of problem solving in technology-rich environment tasks selected for inclusion in the Main Study (approximately eight scenarios of varying lengths). The scenarios finally selected for the Main Study were organized into a pair of 25-minute blocks.

Reading component measures also were constructed according to the framework developed by the literacy expert group. These measures focused on speed and accuracy and were assessed in a limited amount of time. A total of 20 minutes was allotted for the Main Study to measure several of these skills with final measures assembled from 40 minutes worth of Field Test data.

1.5 Main Study goals and design

The assessment design for the Main Study served two primary goals, to: 1) provide good measurement of all the domains included in PIAAC and 2) provide a baseline for assessing trends or changes over time in future rounds of PIAAC or similar assessments. The PIAAC assessment design for the Main Study was based on an assumption of 60 minutes of testing time, on average, for the direct assessment. As PIAAC was not a timed assessment, some respondents were expected to take longer to complete the survey.

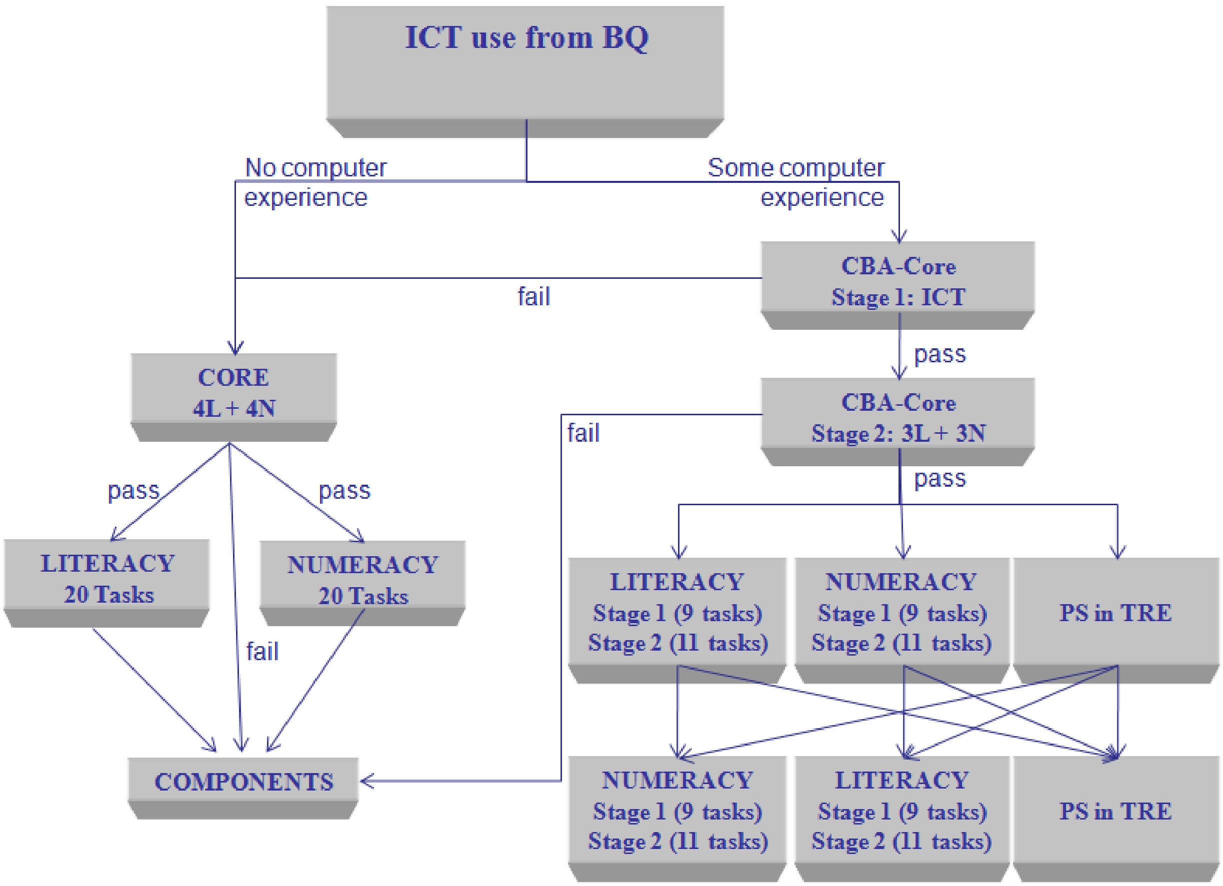
The Main Study design was implemented using the design illustrated below, where L represents literacy tasks, N represents numeracy tasks and PSTRE represents tasks involving problem solving in technology-rich environments. Among other things, the BQ asked about the respondent's computer experiences, which was essential to branch respondents to either the paper-and-pencil or CBAs at the conclusion of the BQ. Respondents with no computer experience based on BQ questions G_04 and/or the H_04a were routed to the paper branch, as were respondents refusing to take the test on the computer. The remainder of respondents were routed to the computer branch of the survey.

As shown in the figure above, the Main Study had the following characteristics:

- The paper-delivered branch of the assessment included a 10-minute core assessment of literacy and numeracy skills. Respondents who performed at or above a minimum standard were randomly assigned to a 30-minute cluster of literacy or numeracy items, followed by a 20-minute assessment of component skills. The relatively small proportion of respondents who performed poorly on the paper-and-pencil core tasks skipped the literacy and numeracy items and were routed directly to the reading component skills measures.

- The computer-delivered branch of the assessment first directed respondents to the CBA Core section, which was composed of two stages taking approximately five minutes each. Poor performance on either stage of the computer-based CBA Core section resulted in switching over to the appropriate sections of the paper-and-pencil instruments. Respondents who failed CBA Core Stage 1 (which contained ICT related tasks) were directed to begin the paper-based Core section and proceed with the process outlined in the above bullet. Respondents who passed CBA Core Stage 1 but failed CBA Core Stage 2 (which contains six cognitive items) were then administered only the reading components tasks. Respondents who performed well on the both CBA Core sections were routed to one of three possible outcomes (each taking approximately 50 minutes): 50% of respondents received a combination of literacy and numeracy tasks, 33% received problem solving combined with either literacy or numeracy, and 17% received only problem-solving sections.

Figure 1.4: Integrated Main Study assessment design



It is also important to note that PIAAC was the first international comparative survey to include multistage adaptive testing as part of the Main Study. The Main Study CBA for literacy and numeracy, represented by each numeracy or literacy block in Figure 1.4, was organized according to the design shown here in Table 1.2. As noted here, the literacy and numeracy modules each consisted of two stages. Each stage contained a number of testlets varying in difficulty. In each stage, only one testlet was delivered to a respondent. Within each of these modules, a respondent took 20 items (nine items in Stage 1; 11 in Stage 2). Thus, respondents taking literacy in Module 1 and numeracy in Module 2 (or vice versa) answered 40 items. Each module was designed to take an average of 30 minutes.

Problem solving in technology-rich environments (PSTRE) is unique because of the nature of the domain. It was organized as two fixed sets of tasks: seven in Module 1 and seven in Module 2. These were also designed to take an average of 30 minutes.

Table 1.2: Design of the Main Study CBA instruments for literacy and numeracy in the integrated design

STAGE 1 (18 unique tasks – 9 tasks per testlet. Each respondent takes 1 testlet)							
	Block A1	**Block B1**	**Block C1**	**Block D1**			
Testlet 1-1	4 tasks	5 tasks					
Testlet 1-2		"	4 tasks				
Testlet 1-3			"	5 tasks			
STAGE 2 (31 unique tasks – 11 tasks per testlet. Each respondent takes 1 testlet)							
	Block A2	**Block B2**	**Block C2**	**Block D2**	**Block E2**	**Block F2**	**Block G2**
Testlet 2-1	6 tasks	5 tasks					
Testlet 2-2		"	3 tasks	3 tasks			
Testlet 2-3				"	3 tasks	5 tasks	
Testlet 2-4						"	6 tasks

However, due to the diversity of the participants' country, language, and educational backgrounds, a deterministic assignment of stages would likely have resulted in certain subpopulations being exposed to only a small percentage of items created for the assessment. To help mitigate the potential impact of such a situation, a set of conditional probability tables of item exposure rates for specified subpopulations was developed. By adjusting these parameter values, a balance between the adaptiveness of the assessment and the predetermined item exposure rates for the given subpopulations was achieved.

Choice of first module: For the computer branch, the selection of a domain (literacy, numeracy or problem solving) for the first module was random. The choice was determined by a random number between 0 and 1 that was generated by the system. A literacy module was chosen if the random number was less than 0.3333333, a numeracy module was chosen if the number was equal to or greater than 0.3333333 and less than 0.6666666, and a problem-solving module if the random number was equal to or greater than 0.6666666.

In problem solving, all respondents took a problem-solving orientation followed by the same set of tasks. In literacy and numeracy, because of the adaptive design, respondents also received the associated orientation but were then assigned to one of the three testlets in Stage 1.

Choice of Stage 1 testlet within literacy and numeracy: The literacy and numeracy testlets in Stage 1 varied in difficulty. There were three levels of testlets: easy (Testlet 1), medium (Testlet 2) and difficult (Testlet 3). Three variables determined which testlet was chosen for a respondent:

- Education level (EdLevel3) from the BQ: Levels were low, medium or high
- Native versus nonnative speaker: The respondent was considered a native speaker if his or her first language was one of the assessment languages
- CBA-Core Stage 2 score: Passing scores between 3 and 6

These three variables were organized in a matrix that results in two threshold numbers. The following matrix provides an example, using Stage 1 selection as explained below in Table 1.3.

Table 1.3: Example of matrix design for Stage 1 selection of literacy and numeracy testlets

EdLevel3:		**Low**		**Low**		**Medium**		**Medium**		**High**	
Native Speaker:		**No**		**Yes**		**No**		**Yes**		**Both**	
Threshold:		I	II	I	II	I	II	I	II	I	II
CBA-Core Stage 2 Score	0	0.900	0.950	0.872	0.922	0.850	0.900	0.822	0.872	0.800	0.850
	1	0.738	0.945	0.710	0.917	0.688	0.895	0.660	0.867	0.638	0.845
	2	0.607	0.924	0.579	0.896	0.557	0.874	0.529	0.846	0.507	0.824
	3	0.505	0.887	0.477	0.859	0.455	0.837	0.427	0.809	0.405	0.787
	4	0.433	0.834	0.405	0.806	0.383	0.784	0.355	0.756	0.333	0.734
	5	0.392	0.765	0.364	0.737	0.342	0.715	0.314	0.687	0.292	0.665
	6	0.380	0.680	0.352	0.652	0.330	0.630	0.302	0.602	0.280	0.580

As shown in the matrix above, if a respondent had a high education level, was a native speaker, and scored high on the CBA-Core Stage 2 (for a total score of 6), he or she would be assigned 0.280 and 0.580 as thresholds. Then a random number between 0 and 1 was generated. This respondent received the easier testlet if the random number was less than 0.280; the medium test if equal to or greater than 0.280 and less than 0.580; and the difficult test if equal to or greater than 0.580. This process ensured that respondents who were native speakers, highly educated, and performed well on the core were most likely to receive the most difficult testlet at the first stage compared to other testlets. However, there was some probability they would receive one of the other easier testlets.

Choice of second testlet for literacy and numeracy module (1): The four literacy and numeracy testlets in Stage 2 also varied in difficulty, with Testlet 1 being the easiest and Testlet 4 the most difficult. For this scenario, three thresholds were defined because there was one more category than in Stage 1. Thus, the test assignment for Stage 2 depended on the following three variables as shown in Table 1.4:

Table 1.4: Example of matrix design for Stage 2 selection of literacy and numeracy testlets

EdLevel3:		Low			Low			Medium			Medium			High		
Native Speaker:		No			Yes			No			Yes			Both		
Threshold:		I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
CBA-Core Stage 2 + Testlet 1 Score	0	0.800	0.900	1.000	0.775	0.875	0.975	0.750	0.850	0.950	0.725	0.825	0.925	0.700	0.800	0.900
	1	0.735	0.871	0.998	0.710	0.846	0.973	0.685	0.821	0.948	0.660	0.796	0.923	0.635	0.771	0.898
	2	0.673	0.841	0.993	0.648	0.816	0.968	0.623	0.791	0.943	0.598	0.766	0.918	0.573	0.741	0.893
	3	0.616	0.812	0.986	0.591	0.787	0.961	0.566	0.762	0.936	0.541	0.737	0.911	0.516	0.712	0.886
	4	0.563	0.783	0.977	0.538	0.758	0.952	0.513	0.733	0.927	0.488	0.708	0.902	0.463	0.683	0.877
	5	0.513	0.753	0.965	0.488	0.728	0.940	0.463	0.703	0.915	0.438	0.678	0.890	0.413	0.653	0.865
	6	0.468	0.724	0.951	0.443	0.699	0.926	0.418	0.674	0.901	0.393	0.649	0.876	0.368	0.624	0.851
	7	0.427	0.695	0.934	0.402	0.670	0.909	0.377	0.645	0.884	0.352	0.620	0.859	0.327	0.595	0.834
	8	0.389	0.665	0.915	0.364	0.640	0.890	0.339	0.615	0.865	0.314	0.590	0.840	0.289	0.565	0.815
	9	0.356	0.636	0.894	0.331	0.611	0.869	0.306	0.586	0.844	0.281	0.561	0.819	0.256	0.536	0.794
	10	0.327	0.607	0.870	0.302	0.582	0.845	0.277	0.557	0.820	0.252	0.532	0.795	0.227	0.507	0.770
	11	0.301	0.577	0.844	0.276	0.552	0.819	0.251	0.527	0.794	0.226	0.502	0.769	0.201	0.477	0.744
	12	0.280	0.548	0.815	0.255	0.523	0.790	0.230	0.498	0.765	0.205	0.473	0.740	0.180	0.448	0.715
	13	0.263	0.519	0.784	0.238	0.494	0.759	0.213	0.469	0.734	0.188	0.444	0.709	0.163	0.419	0.684
	14	0.249	0.489	0.751	0.224	0.464	0.726	0.199	0.439	0.701	0.174	0.414	0.676	0.149	0.389	0.651
	15	0.240	0.460	0.715	0.215	0.435	0.690	0.190	0.410	0.665	0.165	0.385	0.640	0.140	0.360	0.615

- Education level (EdLevel3) from the BQ: Levels were low, medium or high
- Native versus nonnative speaker: The respondent was considered a native speaker if his or her first language was one of the assessment languages

- CBA-Core Stage 2 score plus Stage 1 score: CBA-Core Stage 2 passing scores were between 3 and 6 while the results of Stage 1 were between 0 and 9

These three variables are also organized in a matrix that resulted in three threshold numbers (see matrix below as an example). However, there are now three different matrices, depending on which testlet (easy, medium or difficult) the respondent came from in Stage 1. The appropriate matrix was chosen and the variables were compared with the matrix. This resulted in three threshold numbers for the respondent.

Again, if a respondent had a high education level, was a native speaker, and scored high on the CBA-Core Stage 2 (for example a total score of 6) and had the highest score in Stage 1 (a 9), he or she would be assigned thresholds of 0.140, 0.360 and 0.615. Then a random number between 0 and 1 was generated. Thus, this respondent would have received Testlet 1 (easiest) if the random number was less than 0.140, Testlet 2 if equal to or greater than 0.140 and less than 0.360, Testlet 3 if equal to or greater than 0.360 and less than 0.615, or Testlet 4 (most difficult) if equal to or greater than 0.615.

Choice of second module: After completing Module 1 (either the two testlets for literacy or numeracy or the problem-solving module), the respondent proceeded to Module 2. The selection between Module 1 and Module 2 was also based on random probabilities. Thus, a random number between 0 and 1 was generated again.

- If the respondent completed Literacy as Module 1, he or she was assigned Numeracy as Module 2 (starting with numeracy orientation) if the random number was less than 0.75. Otherwise he or she continued with Problem Solving as Module 2 (starting with PS orientation).

- If the respondent completed Numeracy as Module 1, he or she was assigned Literacy as Module 2 (starting with literacy orientation) if the random number was less than 0.75. Otherwise he or she continued with Problem Solving as Module 2 (starting with PS orientation).

- If the respondent completed Problem Solving as Module 1, he or she was assigned Literacy Module 2 (starting with the literacy orientation) if the random number was less than 0.25, Numeracy Module 2 (starting with the numeracy orientation) if the random number was equal to or greater than 0.25 but less than 0.50, or Problem Solving Module 2 if the random number was equal to or greater than 0.50 (without the PS orientation, which he or she would have already received in Module 1).

After completing the paper or computer branches, the interview continued to the Exit Module, where the interviewer thanked the respondent for participating and provided an incentive, if applicable. The interviewer then continued to the case finalization by answering a set of general questions about the circumstances under which the interview took place, called ZZ-questions.

1.6 Summary and conclusions

This document describes and illustrates the goals and assessment design for both the Field Test and Main Study. The multiple goals of the Field Test illustrate its importance in successfully implementing the Main Study. It was intended to address help evaluate four key areas – operational, platform, instrumentation and scaling and psychometric characteristics. The fact that the results of PIAAC were being linked to previous assessments while being implemented in both paper and computer mode – while also including multistage adaptive testing – added to the importance of the Field Test. Information generated during the Field Test was used to help establish the adaptive portion of the Main Study.

The integrated design included the four cognitive domains as specified in the original terms of reference. As the OECD and the participating countries identified reading components and problem solving in technology-rich environments as international options, alternative designs were also illustrated and described in this chapter. Within the four domains and two formats of PIAAC, the described designs brought innovative aspects and important benefits to the overall goal of producing outcomes that are both valid and comparable across countries.

The Field Test data were used to not only evaluate the procedures and quality of the platform and instruments but to serve to establish the feasibility of linking over time and across modes. The design and data from the Main Study not only expands the range of what can be measured in adult surveys but also how they are measured. More importantly, this information in combination with that gained from the BQ and JRA module described elsewhere in this report provides policymakers and others with a rich source of information to understand the distributions of human capital in their country and the connections between these skills and important social, educational and labor market outcomes. The information from the Main Study was also used to adjudicate the quality of each country's data. This information was shared with the OECD Secretariat, the Board of Participating Countries and all National Project Managers.

ANNEX A1. PAPER-AND-PENCIL INSTRUMENTS

Field test, paper-based instruments in the integrated design where P1-P4 present paper booklets, CL represent the core literacy cluster, CN represent the core numeracy cluster, L1-L2 represent literacy clusters, and N1-N2 represent numeracy clusters.

	Clusters			
Paper-based instruments	**Core (10 minutes)**	**1 (20 minutes)**	**2 (20 minutes)**	**3 (10 minutes)**
P1	CL (6 Lit tasks)	L1 (15 Lit tasks)	L2 (14 Lit tasks)	Components A
P2	CL (6 Lit tasks)	L2 (14 Lit tasks)	L1 (15 Lit tasks)	Components B
P3	CN (6 Num tasks)	N1 (15 Num tasks)	N2 (14 Num tasks)	Components C
P4	CN (6 Num tasks)	N2 (14 Num tasks)	N1 (15 Num tasks)	Components D

ANNEX A2. COMPUTER-BASED INSTRUMENTS

Field test, computer-based instruments with the assessment of reading components where C1-C21 represent computer booklets, L1-L4 represent literacy clusters, N1-N4 represent numeracy clusters, and PS1-PS5 represent problem solving clusters

Computer-based instruments	**Cluster 1 (30 min)**	**Cluster 2 (30 min)**
C1	L1	L2
C2	L2	L3
C3	L3	L4
C4	L4	L1
C5	N1	N2
C6	N2	N3
C7	N3	N4
C8	N4	N1
C9	L1	PS1
C10	L2	PS2
C11	L3	PS3
C12	L4	PS4
C13	N1	PS2
C14	N2	PS3
C15	N3	PS4
C16	N4	PS5
C17	PS1	PS2
C18	PS2	PS3
C19	PS3	PS4
C20	PS4	PS5
C21	PS5	PS1

Chapter 2: The Development of the PIAAC Cognitive Instruments

Mary Louise Lennon and Claudia Tamassia, ETS

2.1 Introduction

As the first computer-based, large-scale assessment of adult skills, PIAAC was designed to reflect the changing nature of information, its role in society and its impact on people's lives. As a result, the cognitive instruments developed for PIAAC differed from those in earlier adult assessments in several important ways.

- For the first time, this assessment addressed literacy in digital environments. As a computer-based assessment, PIAAC was able to include tasks that required respondents to use electronic texts including Web pages, emails and discussion boards. These stimulus materials included hypertext and multiple screens of information and simulated real-life literacy demands presented by digital media.

- The definition of numeracy in PIAAC was broadened from that used in earlier assessments and included the ability to access, use, interpret, and communicate mathematical information and ideas in order to engage in and manage the mathematical demands of a range of situations in adult life. The inclusion of "engage" in the definition signaled that not only cognitive skills but also dispositional elements, that is, beliefs and attitudes, are necessary to effectively meet the demands of numeracy in everyday life.

- PIAAC also included a new domain: problem solving in technology-rich environments (PSTRE). This was the first attempt to assess such a construct on a large scale and as a single dimension. PSTRE included computer-based simulation tasks designed to measure the ability to analyze various requirements of a task, define goals and plans, and monitor progress until task purposes were achieved. The focus was not on computer skills per se, but rather on the cognitive skills required to access and make use of computer-based information to solve problems.

- Finally, PIAAC included a reading components domain, which included measures of vocabulary knowledge, sentence processing and passage comprehension. The inclusion of this domain provided more information about the skills of individuals with low levels of literacy proficiency than had been available from previous international assessments. This was important because to have a full picture of literacy in any society, it is necessary to have information about adults with lower skill levels as it is these individuals who are at greatest risk of negative social, economic and labor market outcomes.

While PIAAC introduced significant new elements to the assessment of adult skills in an international context, key aspects of previous surveys were employed as well. In particular, like the earlier assessments to which PIAAC was linked, this development work was based on frameworks that defined the assessment constructs for each domain as well as features of the tasks designed to measure those constructs.

2.2 Defining the domains: The PIAAC cognitive frameworks

The frameworks for each of the three cognitive domains – literacy (including reading components), numeracy, and problem solving in technology-rich environments – were developed using the same process and methodology. Following Messick's (1994) construct-centered approach, the expert group for each domain defined the construct to be measured, the performances or behaviors expected to reveal that construct, and the task characteristics to be used in building assessment tasks to elicit those behaviors. The overall goal of this process, which included the steps described below, was to explicitly lay out the inferences and assumptions about what was to be measured and how the results would be interpreted and reported.

1. Defining the domain

> Each expert group began by developing a working definition of the domain and the assumptions underlying that definition. Such a definition is an important step in developing an assessment framework as it sets the boundaries for what will and will not be measured.

2. Organizing the domain

> Once the definition was developed, the experts described the kinds of tasks that represent the skills and abilities included under that definition. Those tasks were then categorized to inform test design and, ultimately, score reporting. The goal of this step was to develop a coherent representation of the domain that would permit policymakers and others to summarize and report information in useful ways.

3. Identifying task characteristics

> Step 3 involved identifying a set of key characteristics, or task models, that formed the basis for constructing the assessment tasks. These models defined characteristics of the stimulus materials to be used as well as characteristics of the tasks presented to respondents. Examples of task characteristics used in PIAAC include contexts, material or text types, and task types, which include the cognitive processes or strategies required to complete a given task.

4. Identifying and operationalizing variables

> In order to use the task characteristics in designing the assessment and, later, in interpreting the results, the variables associated with each task characteristic needed to be defined. These definitions are typically based on existing literature and on experience with building and conducting other large-scale assessments.

> This information allowed item developers to categorize stimulus materials as well as the items they constructed so they could be used in reporting results.
>
> As an example, the literacy framework provided further definition of three key task characteristics in that domain: context, text and task type. “Contexts” were defined to include work and occupation, personal uses (home and family, health and safety, etc.), community and citizenship, and education and training. The expert group specified that “texts” could be classified according to medium (print or digital), format, and text type (description, narration, exposition, etc.), and “task types” were defined to include tasks that required respondents to access and identify information, integrate and interpret texts, and evaluate and reflect on information.

Additional steps that follow the Main Study data collection include work to validate the variables that were used to develop the assessment tasks. This includes data analysis to determine which of the variables account for large percentages of the variance in the distribution of tasks and thereby contribute most towards understanding task difficulty and predicting performance. The goal of this analysis is to provide empirical evidence that a set of variables can be identified that summarizes some of the skills and strategies that are involved in accomplishing various kinds of tasks. Finally, an interpretative scheme is built that uses the validated variables to explain task difficulty and examinee performance. The definition of the proficiency levels for each scale, described in greater detail in Chapter 22, is an example of such an interpretative scheme. For previous large-scale literacy assessments, including IALS and ALL, developing these interpretations has provided a useful means for exploring the progression of information-processing demands across each of the scales and for defining what scores along a particular scale mean. In this way, the interpretative scheme contributes to the construct validity of inferences based on scores from the measure on which it is based (Messick, 1989).

The following sections summarize key aspects of the frameworks for the cognitive domains assessed in PIAAC: literacy, reading components, numeracy and problem solving in technology rich environments. The complete framework documents can be accessed at the OECD site at http://www.oecd.org/site/piaac/publications.htm.

2.2.1 Literacy

2.2.1.1 Definition of the domain

In PIAAC, literacy was defined as understanding, evaluating, using and engaging with written texts to participate in society, to achieve one’s goals, and to develop one’s knowledge and potential.

2.2.1.2 Categorizing texts (task characteristics)

A number of variables were used to categorize texts in the PIAAC literacy assessment, including the following:

- **Medium**
  Texts were distinguished as either digital (electronic) texts or print texts. A text that could be reproduced in print exactly as it appears on a screen was considered to be a *print* text.

That is, merely being displayed on a computer screen was not a sufficient condition for classification as a digital text. Texts that could not be reproduced in print with all of their features intact were considered *digital* texts.

- **Format**
  Texts were also classified as either continuous or noncontinuous, with those containing both elements classified as "mixed." Continuous texts are made up of sentences formed into paragraphs. Examples include newspaper articles, brochures, manuals, email and many Web pages. Noncontinuous texts, or matrix documents, include tables, graphs, charts and forms.
- **Type**
  Text types (rhetorical stances) constitute ways of organizing continuous texts in terms of their content and the purpose of the author. Six types of rhetorical stances were identified for PIAAC including: description, narration, exposition, argumentation, instruction and records.
- **Social context**
  The context in which reading takes place may influence the motivation to read and the manner in which texts are interpreted. Therefore, the expert group specified that stimulus materials for the assessment should be drawn from a range of contexts, including: work and occupation, personal (home and family, health and safety, consumer economics, and leisure and recreation), community and citizenship, and education and training

2.2.1.3 Aspects of tasks

Literacy tasks in the PIAAC assessment were designed to address three broad cognitive strategies identified as necessary for achieving a full understanding of texts:

- *access and identify tasks* require respondents to locate information in a text,
- *integrate and interpret tasks* involve relating parts of one or more texts to each other, and
- *evaluate and reflect tasks* require the respondent to draw on knowledge, ideas or values external to the text to evaluate aspects including accuracy, reliability and timeliness.

2.2.1.4 Factors that affect task difficulty

Finally, the Literacy Expert Group defined a number of key factors for item developers to keep in mind as tasks were developed along the continuum from easier to harder.

- **Transparency of information**
  One factor affecting task difficulty is the transparency of information in the text as it relates to the presented task or question. A question that explicitly refers to literal information in a text is generally easier to process and therefore tends to be an easier task along the Literacy scale.

- **Degree of complexity in making inferences**
  Complexity of inferences can be impacted by the extent to which respondents need to recognize paraphrased information, make high-level text inferences, and employ extra-textual inferences.

- **Semantic and syntactic complexity**
  Tasks requiring the reader to identify concrete information such as persons, things or places tend to be easier than those involving abstract properties, such as goals, conditions and purposes. The grammatical structure of the question posed or the stimulus text can also make a task more or less complex. For instance, negative phrases are more complex than affirmative phrases. The presence of subordinate clauses is an example of another feature that can increase the complexity of syntactic processing.

- **Amount of information needed**
  The amount of text that must be processed plays a role in the difficulty of any task. The more information a respondent needs from the text to complete the task, the more difficult that task will be.

- **Prominence of the information**
  Task difficulty can also be impacted by the location of relevant information in a text. It is easier to access information in a prominent location such as in the first or last sentence of a paragraph, in a main, rather than subordinate, clause, or at the top or bottom of a list.

- **Competing information**
  Task difficulty can be impacted by the amount of potentially relevant information the reader has to sift through to access information needed to complete that task. For example, if a text includes telephone, fax and mobile numbers, it will be more difficult for the reader to find the fax number than if the text includes only the fax number.

- **Text features**
  The degree to which the reader has to construct relationships among parts of the text affects difficulty. For example, tasks that require respondents to sort out anaphoric references or which include text where cohesion signals are absent tend to be more difficult.

2.2.1.5 Item development goals for literacy

As part of its work, the Literacy Expert Group was asked to define overall item development targets across the three defined task characteristics of text type, context, and process. For text type, the goal was that 70-80% of the items would be based on print texts and 20-30% on digital texts. The higher percentage of print texts was dictated in large measure by the number of linking items required by the PIAAC assessment design, as those items were developed for paper-based assessments. Both the print and digital categories included continuous and noncontinuous texts.

To ensure a range of contexts in the assessment tasks, the overall targets were to have 15% of items in the work context, 40% in personal, 30% in community, and 15% in education. In terms of task aspects, the framework goals included 40% of the items in the access and identify category, 45% in integrate and interpret, and 15% in evaluate and reflect.

Reading components

In previous assessments of adult literacy, the information gathered on the reading abilities of adults with poor skills was often insufficient to gain a proper understanding of their difficulties due to the small number of items at low difficulty levels. To redress this problem, the literacy framework for PIAAC included a component test intended to provide more information about the abilities of those with low levels of literacy.

The components assessment framework was based on the principle that comprehension – the process of constructing meaning when reading – is built on knowledge of how a given language is represented in its writing system and through component print-reading skills. Evidence of an individual's level of print-reading skills can be captured in tasks that examine a reader's ability and efficiency in processing the elements of the written language, including letters/characters, words, sentences, and larger, continuous segments of text.

A second guiding principle is that the assessment of component skills aims to evaluate the extent to which adults can apply their existing language and comprehension skills to the processing of printed texts. The components tasks were not designed to separately assess the level of language skills in the target writing system and the literacy skills assessed in the main literacy survey. Nonnative speakers of the language of the assessment who have only basic oral vocabulary, syntactic/grammatical and linguistic comprehension skills were expected to show poor performance on component reading tasks. As a consequence, low levels of proficiency in the language of the assessment were not differentiated from low literacy skills in the component tasks.

A third guiding principle is that the levels of proficiency, efficiency and integration of component skills are indicative of the levels of reading development and learning potential. As skills and knowledge accumulate, the ease of processing familiar, text-based print increases. Component efficiency is typically indexed by assessing speed or rate of processing, as well as accuracy. For PIAAC, although the reading components assessment was the one domain assessed only in paper-and-pencil form, interviewers timed respondents and recorded that information as part of the measure of efficiency.

It was also assumed that the set of component items administered in each country reflected the linguistic characteristics of the language of assessment. As the relationship of the language to the writing system was anticipated to be very different in different languages, the nature of the items used to assess the components was adapted based on consideration of those differences in order to best ensure comparability across languages. Countries were provided with very specific adaptation guidelines and training on how to adapt the reading components measures for their language(s) of assessment. As was true for the other domains, trained verifiers reviewed these adaptations and provided feedback to countries as needed.

The PIAAC components assessment included tests of vocabulary, sentence processing, and basic passage comprehension. In skilled reading, these components are integrated to support literacy performance. During acquisition, even by adults, these components may be measured separately, with different profiles having implications for learning, instruction, and policy.

2.2.2 Numeracy

2.2.2.1 Definition of the domain

PIAAC defined numeracy as *the ability to access, use, interpret and communicate mathematical information and ideas, in order to engage in and manage the mathematical demands of a range of situations in adult life.*

Because numeracy is a broad, multifaceted construct referring to a complex competency, the definition of numeracy was coupled with a more detailed definition of *numerate behavior* and with further specification of the facets of numerate behavior. The expert group felt this was necessary for the operationalization of the construct of numeracy in PIAAC and to broaden the understanding of key terms appearing in the definition itself. The definition of numerate behavior adopted for PIAAC was as follows, with key facets or task characteristics associated with numerate behavior shown in Table 2.1.

Numerate behavior involves managing a situation or solving a problem in a real context, by responding to mathematical content/information/ideas represented in multiple ways.

Table 2.1: Numerate behavior – key facets and their components

Numerate behavior involves managing a situation or solving a problem…
1. in a real context:
– everyday life
– work
– society
– further learning
2. by responding:
– identify, locate or access
– act upon and use: order, count, estimate, compute, measure, model
– interpret
– evaluate/analyze
– communicate
3. to mathematical content/information/ideas:
– quantity and number
– dimension and shape
– pattern, relationships, change
– data and chance
4. represented in multiple ways:
– objects and pictures
– numbers and mathematical symbols
– formulae
– diagrams and maps, graphs, tables
– texts
– technology-based displays
5. Numerate behavior is founded on the activation of several enabling factors and processes:
– mathematical knowledge and conceptual understanding
– adaptive reasoning and mathematical problem-solving skills
– literacy skills
– beliefs and attitudes
– numeracy-related practices and experience
– context/world knowledge

2.2.2.2 Principles for assessing numeracy in PIAAC

The development of the numeracy assessment for PIAAC was based on a number of general principles or guidelines, as listed below:

- **Items should cover as many aspects as possible within each of the four facets of the numeracy competency.**
  Items should require the activation of a broad range of skills and knowledge included in the construct of numeracy.

- **Items should aspire to maximal authenticity and cultural appropriateness.**
  Tasks should be derived from real-life stimuli and pertain to a range of contexts or situations (i.e., everyday life, work, society, further learning) that can be expected to be of importance or relevance in the countries participating in PIAAC. Item content and questions should appear purposeful to respondents across cultures.

- **Items should have a free-response format, to the extent feasible within the computer platform used for administering the direct assessments.**
  Items should be structured to include a stimulus (e.g., a picture, drawing, visual display) and one or more questions, the answers to which the respondent communicates via the modes available within the test platform, primarily: numeric entry, click, highlight a region of the stimulus, or use of various pull-down menus.

- **Items should spread over different levels of ability**
  Items should span the range of ability levels anticipated among PIAAC participants, from low-skilled individuals to those with advanced competencies.

- **Items should represent the different response types**
  Items should require the range of available response types. It was recognized that certain types of numeracy responses, especially those requiring the use of interpretation, evaluation, analysis and communication, could receive only partial coverage in the first cycle of PIAAC due to the constraints of automatic scoring.

- **Items should vary in the degree to which the task is embedded in text**
  Some items should use relatively rich texts while others should use little or no text. This distribution aims to reflect the different levels of text involvement in real-world numeracy tasks, as well as minimize overlap with the literacy assessment.

- **Items should be efficient**
  To allow for coverage of many key facets of the numeracy competency, a large number of diverse stimuli and questions should be included. However, given testing-time constraints, the use of short tasks is necessary, precluding items that can simulate extended problem-solving processes or require a lengthy open-ended response.

- **Items should be adaptable to unit systems across participating countries**
  Items should be designed in a way that their underlying mathematical demands are as consistent as possible across countries, regardless of language and mathematical conventions. After being translated, items should retain equivalency with respect to their mathematical or cognitive demands.

2.2.2.3 Item development goals for numeracy

As was the case for literacy, part of the development work for the expert group included defining item development goals across the key facets of numeracy as defined in the framework. For response, or process, facets the goals included 50% of items in the act upon and use category, 10% in identify, locate or access, and 40% in interpret and evaluate. The framework specified that tasks should be based on real-life stimuli appropriate to a range of contexts or situations (i.e., everyday life, work, societal, further learning) without outlining specific proportions in each category. For mathematical content, development goals included a distribution of 25% of the items relating to data and chance, 25% dimension and shape, 20% pattern, relationships and change, and 30% quantity and change.

2.2.3 Problem solving in technology-rich environments (PSTRE)

2.2.3.1 Definition of the domain

PSTRE was broadly defined as *using digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks.*

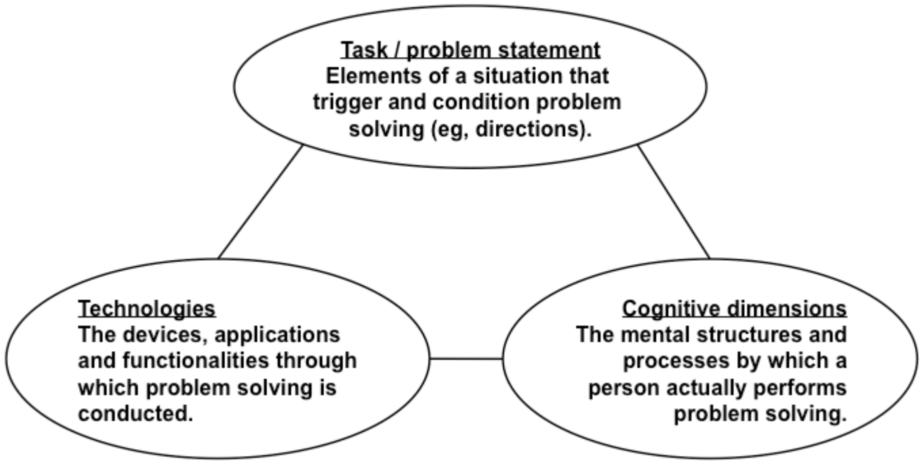
PIAAC represented the first attempt to assess PSTRE on a large scale and as a single dimension. This presented challenges in terms of the definition of tasks and the practical collection of data. Furthermore, digital technologies continue to evolve at a rapid pace, as do the personal, social and work-related uses of these technologies. While setting the stage for further rounds of assessment, the framework took into consideration issues of feasibility as well as the evolution of technology and its uses. In light of these challenges and constraints, the definition went on to further specify the scope of this first assessment of PSTRE for PIAAC:

The first PIAAC problem-solving survey focuses on the abilities to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, and accessing and making use of information through computers and computer networks.

2.2.3.2 Core dimensions of problem solving in technology-rich environments

The domain of PSTRE was conceived along three dimensions, as shown in Figure 2.1.

Figure 2.1: Core dimensions of problem solving in technology-rich environments



"Cognitive dimensions" include the mental structures and processes involved when a person solves a problem. These include setting goals and monitoring progress; planning; accessing and evaluating information; and making use of information by selecting, organizing and transforming information.

"Technologies" are the devices, applications and functionalities through which problem solving is conducted. These include hardware devices (laptop computers in the case of PIAAC); simulated software applications; commands and functions; and representations (text, graphics, etc.).

"Tasks" are the circumstances that trigger a person's awareness and understanding of the problem and determine the actions needed to be taken in order to solve the problem. Ordinarily, a wide range of conditions can initiate problem solving. For instance, a computer user may realize that his or her mailbox is crowded and that a new schema is needed for classifying emails. Alternatively, he or she may be faced with a complex issue (such as finding out more about a medical treatment) and decide to look for relevant information on the Web. In test-taking contexts, tasks are more explicitly assigned to respondents. They include the question and task instructions presented to respondents, as well as the specific materials and time constraints associated with the test.

Dimensions of the tasks being assessed in PIAAC PSTRE included:

- *Task purposes and contexts, including personal, work/occupation, and civic*

- *The intrinsic complexity of the problem*
  Intrinsic complexity is related to a set of more specific variables: the minimum number of steps or actions required to solve the problem; the number of options at each phase; the diversity of operators and the complexity of mental reasoning and/or computation; the probability of impasses or unexpected outcomes; the number of constraints to be satisfied; and the amount of composition or transformation needed to communicate a solution.

- *The explicitness of the problem statement and task directions given to the respondent*
  This dimension ranges from well-defined, explicit problem statements to implicit and ill-defined problem statements. A problem situation that requires the respondent to select operators and subgoals or define the successful achievement of a goal makes the problem more difficult.

2.2.3.3 PSTRE in relation to other domains of PIAAC

The constructs of literacy, numeracy and PSTRE rely on the same "core" cognitive processes. For example, tasks in all three domains require both an ability to decode printed symbols and a minimal working memory capacity. PSTRE also assessed a set of competencies distinct from those defined in the other two constructs.

The assessment of PSTRE in PIAAC focused on goal setting, monitoring and planning in technology-rich environments, and assessment tasks emphasized the problem-finding and problem-shaping processes typically found in these environments. Tasks included selecting an appropriate software application; deciding on one among several possible strategies; making use of adequate functionalities in a context-sensitive manner; interpreting ill-structured texts; and using online forms.

Respondents needed to complete problem-solving tasks in environments that involved multiple and complex sources of information. Some of the tasks required respondents to use and shift across multiple environments. PSTRE therefore assessed decision making with respect to the use of information sources (for example, choosing which environment to use or deciding whether or not to go to another website.) Evaluation was included as a critical underlying part of problem solving. Additionally, the selection of appropriate devices or tools took a prominent role in this domain.

In terms of processing information, problem solving is a specific construct in that it focuses on:

- the evaluation of sources in terms of reliability and the adequacy of information relative to the problem statement, as opposed to mere topical relevance, which is more applicable for literacy

- the integration of information across sources, especially in cases where the sources provide inconsistent information

PSTRE tasks sought to minimize the numeracy and literacy demands placed on respondents in order to increase the specificity and validity of the construct.

2.2.3.4 PSTRE and ICT competence

What differentiates the problem-solving domain from the general ICT domain? ICT skills may be broadly defined as "the interest, attitude, and ability of individuals to appropriately use digital technology and communication tools" (Lennon, et al., 2003). As is true for literacy and numeracy skills, ICT skills underlie PSTRE. However, the PSTRE construct aimed to encompass more than the purely instrumental skills related to the knowledge and use of digital technologies.

The cognitive dimensions of problem solving were considered the central object of the assessment, with the use of ICT as secondary.

2.2.3.5 Item development goals for PSTRE

Like literacy and numeracy, the PSTRE Expert Group defined targets for the distribution of items across the categories defined in the PSTRE framework. Based on the development of 25 tasks to be considered for the Field Test, goals included the distribution shown in Table 2.2. Additionally, the distribution across contexts was recommended to be 40% personal, 30% occupational and 30% civic. Finally, the task dimensions of intrinsic complexity and explicitness of the problem definition were specified as development variables as they were expected to influence the difficulty of items in the problem solving assessment.

Table 2.2: Distribution of PSTRE tasks as a function of environment and cognitive dimensions

Cognitive dimensions	Web environment	Spreadsheet environment	Email environment	Multiple environments
Goal setting and monitoring progress	2	1	1	1
Planning	2	2	2	4
Accessing and evaluating information	3	0	0	0
Selecting, organizing and transforming information	2	1	3	1
Totals	9	4	6	6

2.3 Developing the cognitive instruments

2.3.1 Overview

For each of the cognitive domains, test developers worked closely with the expert group to ensure that the instruments reflected the frameworks. All items were also submitted for country review to receive input on cultural and linguistic appropriateness as well as item content. In the case of literacy, developers from Australia and the United States attended each of the expert group meetings and the experts reviewed items throughout the development process. ETS developed the reading components tasks and the Literacy Expert Group reviewed those items as well. For numeracy, the expert group itself assumed primary responsibility for developing the PIAAC items. Test developers reviewed those items to ensure consistency in instructions, response modes and presentation across domains. ETS was primarily responsible for developing the PSTRE tasks and developers met with that expert group to receive input and reviews throughout the development process.

Two core requirements for PIAAC had important implications for development of the cognitive instruments. First, because the domains of literacy and numeracy had been measured in previous large-scale international surveys, it was a requirement that PIAAC link back to the ALL and IALS. As a result, sets of linking items needed to be selected for literacy and numeracy that fit the requirements of the PIAAC assessment design. As described in the following section, transitioning those paper-based linking items to PIAAC's computer-based delivery mode required considerations related to display and response mode issues as part the development process.

A second requirement was that all items be scored by computer. This was a necessary feature in order to implement adaptive testing in PIAAC. Developers thus had to define response modes that could be computer scored across languages for each of the cognitive domains and for both linking and new items in the assessment. The PIAAC design called for the continued use of open-ended response items both to maintain the real-life focus of the assessment and to maintain the psychometric link between PIAAC and prior surveys. While those prior paper-and-pencil surveys allowed respondents to write responses ranging from a word or two to several sentences, the use of automated scoring for such responses was not possible for PIAAC given that the assessment was to be delivered in 33 languages.

The Consortium therefore relied on evidence from previous ETS work on a derivative computer-based test for individuals to define a set of computer-scoreable, open-ended response modes. This work had shown that item parameters for paper-and-pencil items were not impacted when those items were adapted to allow respondents to click on responses, type numeric answers, and highlight responses in text. Development therefore proceeded on the assumption that linking items could be adapted to employ these response modes and still maintain item parameters from previous assessments, an assumption that was ultimately supported by the Field Test data.

Additionally, each of these three response modes required only basic computer skills – an important consideration given that the test needed to be accessible to adults with a range of computer experience. The three are described in more detail below.

- **Clicking items**
  These items required respondents to click on graphical elements, cells in a table, links on a Web page, or radio buttons or check boxes to answer. Respondents could select and change their answers while working within each unit. In terms of scoring, one or more correct responses were defined for each item. This response mode had an advantage in that, in general, click areas remained consistent across languages and therefore scoring did not require much adaptation across different national versions of the items.

- **Numeric entry items**
  For these items, respondents answered by typing a numeric response using the number keys, decimal point (period or comma as appropriate across participating countries) and space key. In this response mode, all other keys on the keyboard were locked and not available for use to prevent respondents from including text in their responses that could not be scored.

  Numeric entry items were scored automatically based on the definition of correct numeric response(s) included in the scoring rule. One scoring rule employed a number match. In this case, a response was correct as long as it represented the correct numerical value, regardless of how that number was represented. For example, if a correct response was 4, responses such as 12/3 or 2*2 would receive a correct score. The second type of numeric scoring rule required an exact match. That is, instead of checking for numerical equivalence, the system checked for character equivalence. In this case, a response of 229 would be scored differently from responses such as 229.0 or 229.00. As described in more detail later in this section, guidelines were provided to allow countries to adapt numbers and number formats in order to present respondents with realistic numerical values in the context of presented tasks.

- **Highlighting items**
  These items allowed respondents to highlight one or more words, phrases and sentences in a text to answer questions. Defining the scoring rubrics for these items was most challenging as responses were language dependent. For each response, developers defined a minimum correct response, as well as a maximum correct response. They based those judgments on ETS's previous work to develop open-ended, computer-scoreable items as well as experience in scoring paper-based responses. In previous paper-based assessments, respondents were given credit for correct answers when they underlined or circled information in the stimulus instead of writing an answer on a response line. Existing rules for what constituted a correct response in those situations thus helped guide the development of rules for highlighted responses in PIAAC. As this was the most language-dependent response mode, countries were actively involved in implementing and testing the minimum/maximum rules for their national versions of these item types.

In terms of the scope of item development for the cognitive instruments, the PIAAC assessment design specified the number of items to be developed for the Field Test and subsequently used in the Main Study. The Field Test and Main Study needs for literacy and numeracy, the two domains with linking items, are shown in Table 2.3. The Main Study design included 24 items for the paper-and-pencil version (19 linking items and five new items) and 48 items for the computer-based version (29 linking and 19 new) for each domain. To reach these goals for the Main Study, the Field Test design specified 35 paper-and-pencil items (25 linking and 10 new) and 72 computer-based items (42 linking and 30 new). Note that for both domains, the Main Study design additionally specified that a set of 18 linking items was to be used in both the paper-and-pencil and computer-based versions of the instruments.

Table 2.3: Literacy and numeracy item needs for PIAAC

	Field Test		**Main Study**	
	Linking	**New**	**Linking**	**New**
Paper Version	25	10	19	5
Computer Version	42	30	29	19

Reading components tasks were developed according the framework for this domain. These measures focused on speed and accuracy and several measures were assessed in a defined amount of time. A total of 20 minutes was allotted in the Main Study to measure these skills, with final measures assembled from 40 minutes worth of Field Test items.

The assessment of PSTRE involved scenarios of varying complexity and length, designed to take between five and 15 minutes to complete. Overall, 14 units were used in the Field Test. Several of those units included multiple parts, or tasks, so a total of 24 tasks were included. Two 25-minute blocks were included in the Main Study. Block 1 had five units, with seven associated tasks, and Block 2 had six units, also with seven tasks.

2.3.1.1 Selecting and adapting linking items

The assessment design for the PIAAC Main Study required that 60 percent of the literacy and numeracy items be taken from, and therefore link back to, previous surveys. In the case of literacy, items from both IALS and ALL were reviewed as potential linking items for PIAAC. As numeracy was not a domain in IALS, all numeracy linking items were selected from the ALL survey. The following aspects were taken into consideration when selecting linking items for inclusion in PIAAC.

- **Item quality**
  To be eligible for inclusion in PIAAC, items needed strong statistics from previous assessments. That is, developers were looking for items with good item parameters and items with no history of differential item functioning or translation problems.

- **Distribution according to the dimensions of the frameworks**
  Items were reviewed and reclassified according to the PIAAC frameworks and, to the extent possible, selected to reflect the distributions recommended by the expert groups.

- **Distribution across levels of difficulty**
  The difficulty of items was taken into consideration in an effort to be sure items reflected the five levels used to report results for both previous studies and PIAAC.

- **Cultural appropriateness**
  Countries were asked to review the selection of linking items to identify any of particular concern in terms of their appropriateness across the range of cultures among PIAAC participating countries.

An additional critical consideration for PIAAC was the suitability of these linking items for computer delivery. All of these items had been developed for paper-and-pencil assessments with open-ended responses that were human scored. For PIAAC, items needed to be computer scored, so selected items needed to be answerable using the response modes of clicking, numeric entry and highlighting. In addition, the stimulus materials for selected items needed to be adaptable to onscreen presentation keeping the same formatting as that used on paper.

The Literacy and Numeracy Expert Groups met in 2008 to review and provide input regarding the selection of linking items for the Field Test as well as to discuss issues associated with moving these items to the computer.

2.3.2 Developing new items

New items were developed to reflect the PIAAC frameworks and take advantage of the computer-based nature of the assessment. For example, new literacy items were designed to assess skills and knowledge associated with digital texts. Literacy and numeracy development also needed to complement the set of items selecting as linking items. As a new domain, PSTRE included only newly developed items. For all domains, the new item development process involved countries, the PIAAC Consortium, and expert groups.

2.3.2.1 National submissions

Countries were invited to participate in the process of developing new items for PIAAC. As is the case with any large-scale international survey, it was important that the pool of tasks for PIAAC reflected the range of contexts and experiences of respondents across participating countries. One way to better ensure this was to solicit national submissions once countries had been introduced to the PIAAC frameworks. The request for literacy and numeracy item submissions was issued in 2008 during the first meeting of the NPMs. The Consortium developed a document that provided: i) a general overview of the item development task, including a description of the scope of work, ii) a summary of the development process to be followed, iii) procedures for submission and review of items, iv) the item development timeline, and v) sample items that illustrated the kinds of items to be developed.

Due to the tight development schedule, countries had three months to develop and submit items. To facilitate country participation, the Consortium accepted item submissions in six languages including English, French, Spanish, German, Japanese and Italian. Additionally, to better integrate submissions into the development process, countries were encouraged to submit items progressively as they were developed, rather than as a single submission close to the deadline.

In preparing materials for submission, national item developers were asked to provide the following information about each item:

- information about the source of the item (original, or from a book or other source)
- information about any copyright considerations for the stimulus materials (who holds the copyright, who had been contacted to seek permission to use the material, and copyright permission when it was obtained). Countries were responsible for obtaining copyright information for any submitted material.
- the classification of each item according to categories in the relevant domain framework

Countries were also encouraged to submit additional stimulus materials without associated items. Wherever possible, the Consortium developed items based on these stimuli in order to ensure a mix of materials that reflected the diversity of cultural contexts represented across participating countries.

Literacy submissions were received from Austria, Estonia, France, Italy and Japan. In Numeracy, submissions were received from Austria, Estonia, Finland, France, Hungary, Italy, Japan and Korea. Submissions received from countries were reviewed and evaluated in terms of their fit to the PIAAC frameworks and contribution to the item pool. This process was documented and summarized in a detailed report that was shared with countries.

Because PSTRE was a new domain with complex development demands, the Consortium did not expect countries to submit fully drafted tasks. Instead, it asked countries were asked to submit ideas for tasks which illustrated common adult uses of technology in problem-solving contexts or where the appropriate use of technological functions (such as a “compare” function on a shopping site or “sort” function in a spreadsheet) facilitated solving a problem. Additionally, countries were encouraged to provide examples of Web sites and other technology environments that they viewed as representative of materials used by adults in their home, community and work environments.

2.3.2.2 Item development

New assessment materials for the Field Test were developed based both on materials submitted by countries and materials developed by the contractors. The development period extended from early 2008, with the first meeting of the expert groups, to early 2009, when the expert groups finalized the selection of the Field Test item pool.

As previously mentioned, in the case of literacy, the PIAAC contractors, including item writers in Australia and the United States, developed the new items. The process differed for numeracy, where the expert group itself drafted all the new items. To accommodate this work, several additional expert group meetings were held. In August 2008, the numeracy expert group met in Dublin, Ireland, and developed approximately 60 items. In November 2008, the group met again in Frankfurt, Germany, to review countries’ comments on the first batch of materials, consider how best to implement the suggested changes, and review other available items. As a new domain, PSTRE also had a higher level of involvement from experts with two additional meetings. The PSTRE Expert Group met in August 2008 in Poitiers, France, to review an initial set of draft tasks. A second meeting, held in Amsterdam in December 2008, included programmers as well as item developers so that features of the simulated technology environments could be discussed and agreed to along with content for specific tasks.

For each domain, stimulus materials were selected based on specifications provided in the framework for that domain. To the extent possible, stimuli for the PIAAC assessment were taken from real-world materials such as newspaper and magazine articles, advertisements, books, forms, and Web pages that adults ages 16-64 would encounter in a range of everyday life contexts. Given the international context of the assessment, care was taken to select materials appropriate across cultures and languages. Soliciting materials from participating countries and having all countries review the stimulus materials were important steps to better ensure this diversity.

It was also important to ensure that stimulus materials would not become too easily dated. Those that contained dates or references to contemporary individuals or events – particularly if such information was central to completing tasks associated with those materials – could become dated by the time the assessment was administered. Such materials were also avoided as they would become increasingly problematic in future testing cycles if they were needed as linking items.

Tasks for PSTRE were situated in simulated computer environments including a browser, email system, spreadsheet and word processor. While these did not replicate the full functionality of real-life environments, they included many key functions. For example, the email environment allowed respondents to reply, reply to all, forward, send and move emails to folders. In the browser environment, respondents could navigate using the back, forward and home buttons and they could bookmark pages for later reference. Presenting the PSTRE tasks in these simulation environments allowed the computer to capture a variety of process information. For any given task, collected information included time spent, actions taken (e.g., clicking and typing responses or selections from drop-down menus such as “file” and “edit”) and the sequence in which actions were completed. This information provided direct evidence of the processes and strategies respondents used to complete assigned tasks and therefore allowed for better inferences about their knowledge and skills related to PSTRE.

2.3.2.3 Item reviews

As an additional step to better ensure that the new items reflected the range of contexts and experiences of respondents across participating countries and to obtain input about item content, all participating countries reviewed the PIAAC item pool at several stages. Guidelines were developed for the review process which specified that the materials were to be reviewed in relation to:

- coding based on the task characteristic categories in the frameworks
- the overall appropriateness of each item. Items were to be classified into one of three categories: acceptable as is, acceptable with modifications, or unacceptable. For the second category, countries were asked to specify revisions that would make the item acceptable. They were also asked to specify the reason or reasons why they rated any items as unacceptable.
- cultural concerns
- translation concerns

Countries were given an opportunity to review draft items before developers finalized them with input from the expert groups. Reviews were conducted in three batches as described below:

- A first batch of new tasks was released on 21 October 2008 with comments due on 7 November. This batch included: i) four item sets for reading components; ii) 16 literacy units with 105 tasks; iii) 20 numeracy units with 48 tasks; and iv) 11 PSTRE scenarios.
- A second batch of new tasks was sent to countries for review on 17 December 2008 with comments due on 20 January 2009. This included a set of 31 new numeracy tasks and six new literacy tasks.
- A third batch was released on 15 January 2009 with comments due on 29 January 2009. This last batch included seven tasks for PSTRE.

2.3.3 Additional supporting materials

The development process for PIAAC cognitive instruments included several sets of materials beyond the items themselves. These included a set of detailed guidelines to assist countries in translating items and scoring guides so that national instruments would remain comparable with the international masters. Equally important was a set of interactive tutorials that introduced respondents to the PIAAC instruments, ensuring that all participants approached the survey with the same information about how to navigate through the assessment and provide their responses.

2.3.3.1 Translation/adaptation guidelines and scoring guides

To support the work of countries in translating and adapting items, implementing computer-based scoring, and translating scoring guides for the paper-based items, the Consortium developed translation and adaptation guidelines as well as master scoring guides for participating countries. These materials also supported the linguistic quality control process, described in Chapter 4, that was designed to help ensure that instruments across countries were comparable and that consistent scoring procedures were implemented.

A sample set of guidelines for one of the Field Test items is shown in Figure 2.2. The guidelines specified linguistic considerations for translation (e.g., maintaining a literal match between wording in the question and stimulus) and defined the correct response for both the paper-and-pencil and computer versions (minimum and maximum) of the item.

Figure 2.2: Sample translation and scoring guidelines

Item Notes: *Translation must maintain literal match between keywords "gym bench" and in question and in table heading under "Muscle building."*

"Muscle" appears in question and four places in the stimulus.

		English Paper and Pencil (same version as ALL)	English Computer
Directions		Use the exercise equipment chart on the opposite page to answer questions x through y.	Look at the exercise equipment chart. Click on the chart to answer the question below.
Question		Which muscles will benefit most if you use the gym bench?	Which muscles will benefit most if you use the gym bench?
Answer	Abdominal (muscles)	<u>Minimum correct response</u>: Clicks on "abdominal muscles" cell <u>Maximum correct response (See illustration below)</u>: Muscle building Gym bench Image of gym bench Very good (intersection of abdominal muscles row and gym bench column) Abdominal muscles	

Maximum correct response

Effects on…	Cardio-Training					Muscle Building							
	Exercise bicycle	Rowing machine	Stepper	Tread-mill	Air trainer	Dumb-bells, weights	Elastic	Gym bench	Muscle-building bench	Multi-trainer	AB trimmer	AB shaper	AB roller
Arm strength	Ineff-ective	Good	Average	Ineff-ective	Good	Very good	Very good	Good	Good	Good	Very good	Good	Good
Leg strength	Good	Very good	Average	Very good	Good	Ineff-ective	Good	Average	Good	Good	Ineff-ective	Good	Good
Abdo-minal muscles	Average	Very good	Good	Good	Average	Ineff-ective	Good	Very good	Good	Average	Very good	Very good	Very good
Overall muscle building	Ineff-ective	Very good	Ineff-ective	Average	Ineff-ective	Average	Good	Good	Good	Average	Good	Good	Good
Heart/ arteries	Very good	Good	Very good	Very good	Good	Ineff-ective	Average	Average	Average	Good	Average	Average	Aver-age
Flexi-bility	Ineff-ective	Good	Ineff-ective	Ineff-ective	Average	Average	Average	Good	Ineff-ective	Ineff-ective	Average	Good	Good
Joints	Good	Very good	Good	Good	Good	Good	Average	Average	Good	Good	Average	Average	Aver-age
Slim-ming	Good	Average	Very good	Good	Good	Ineff-ective	Average	Good	Average	Average	Good	Good	Good
Dangers	None	Back	None	Legs		It is best to learn to use these types of apparatus properly before you make a major effort							

For numeracy, specific guidelines were provided to guide countries in the adaptation of numbers and number formats. For example, two options were provided to address the challenge of consumer-related items that involved currency. The first was for countries to keep the numbers the same but change the currency sign. This was the option of choice for adapting U.S. dollars to euros as the two are close in value. For currencies where simply changing the currency sign would result in unrealistic numbers, a second option was provided. Guidelines specified that in this case, numerical values could be changed by multiplying or dividing them by powers of 10 (and only powers of 10). This restriction was intended to allow countries some flexibility while maintaining similar cognitive demands across national versions.

2.3.3.2 Tutorials

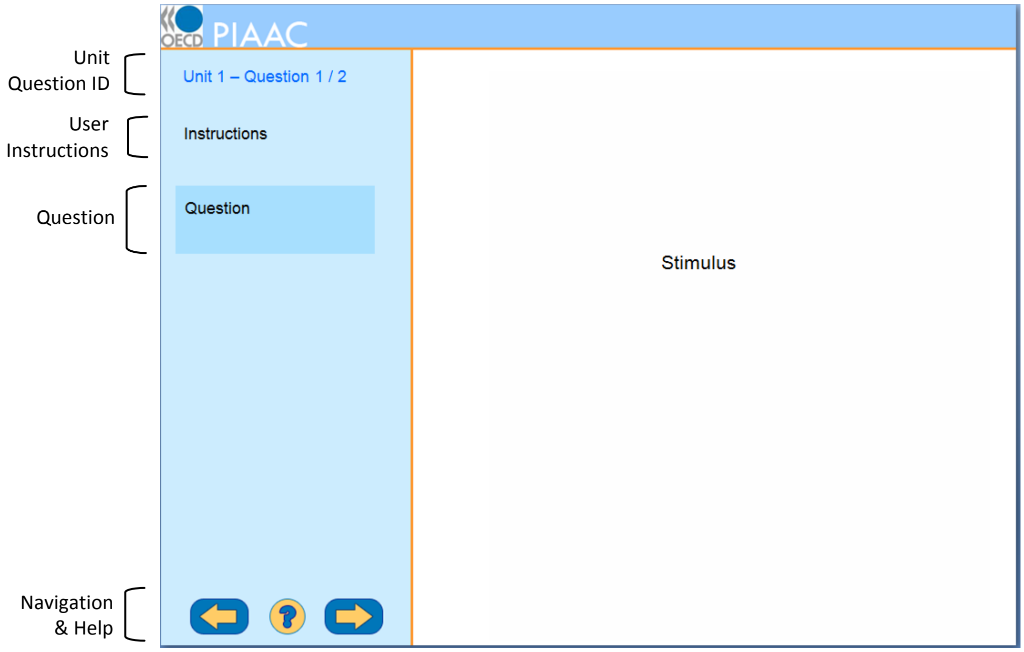
As part of the development process, developers designed a set of tutorials to introduce respondents to the design and layout of the testing screen, familiarize them with the available response modes in each domain, demonstrate the navigation and help functions, and, in the case of PSTRE, define the tools and functionality in the simulated environments. For the Main Study, these tutorials were designed to be relatively brief, about five minutes per domain, in order to reduce respondent burden in terms of the time required to complete the full assessment.

2.3.3.3 PIAAC testing screen

While not a material per se, an additional step in the development of the cognitive instruments was the design of the testing screen for the cognitive items. An important goal was to develop a design which allowed a consistent display and interface across the cognitive domains. PIAAC used a vertically split screen as shown in Figure 2.3. For all domains, the stimulus material was presented on the right and the task information on the left.

Navigation and help icons were located at the lower left. Labels for these icons displayed when the respondent held the cursor over them, allowing translations of various length to display. The user instructions both identified the stimulus and provided information about the required response mode. For example, “Look at the exercise equipment chart. Click on the chart to answer the question below.”

Figure 2.3: PIAAC testing screen



The design presented a number of important advantages.

- The vertical split facilitated the display of paper-and-pencil linking items being moved to the computer. Splitting the screen vertically allowed a display area for stimulus materials that was taller than it was wide. Because this more closely mirrors the width-to-height ratio of paper, this was an advantage for displaying paper-based linking items.
- By not extending the full width of the screen, stimulus text could be formatted with more natural line lengths, improving readability.
- More vertical height accommodated displays across a variety of languages.

2.3.4 Preparation of final Field Test instruments

2.3.4.1 International master

The Consortium finalized and released the master versions of the Field Test items to countries for translation and adaptation according to the timeline shown below. Each round of released items included the items themselves, translation and adaptation guidelines for the items, Verification Follow-up Forms (used for monitoring and documenting the translation/adaptation process), and scoring guides.

- Linking items for numeracy and literacy in computer-based format were released for translation in two rounds: 16 January 2009 and 5 February 2009.
- New numeracy items were released on 6 April 2009. The scoring guides for the paper-based numeracy items were released on 28 April 2009.

- Reading components items were released on 6 April 2009.
- New literacy tasks were released on 9 April 2009. The scoring guides for the paper-based literacy items were released on 28 April 2009.
- PSTRE scenario were released in batches with five scenarios were released on 29 April 2009, five scenarios on 3 May 2009, and four on 30 May 2009.

Master versions of the Field Test paper booklets were also released to countries in the spring of 2009. The assessment design for the Field Test required four sets of paper booklets including: two literacy booklets, two numeracy booklets, and four reading components booklets. The assembly of all paper-based booklets, including instructions for administration and scoring sheets, occurred during this period.

Finally, once the master versions of the computer-based units were tested and finalized, these were assembled into computer blocks following the Field Test design that required four numeracy and four literacy blocks, each with 18 tasks. These blocks were organized in a way that ensured a balanced distribution across important aspects of the frameworks and known or estimated difficulty levels and assembled by the Consortium.

2.3.4.2 National versions

Countries developed their own national versions of the Field Test assessment materials following the translation, adaptation and verification processes developed for PIAAC between April and June 2009. Layout checks were conducted by both the Consortium and countries to identify any display issues requiring modification. Such revisions were prompted by issues including text that did not fit within a table cell due to longer word lengths in some languages, and so on. The Consortium manually fixed layout issues on a case-by-case basis and submitted them to countries for final review and approval.

During this period, countries were also responsible for defining and adapting the computer-based scoring for their national versions where applicable. That is, all language-dependent scoring rules – such as highlighting areas – were defined by the national centers and verified as part of the quality assurance process.

2.3.4.3 Scoring testing

The Consortium tested the automatic scoring for the international version of the literacy and numeracy units prior to distributing the national versions. Two sources of error were observed during international testing: i) errors at the level of item editing, that is, the scoring information was specified incorrectly by the item editor (specification error), and, ii) errors at the level of technology, that is, the software did not work accurately (implementation error). All detected errors were fixed, and the scoring procedures of affected units were retested until no further errors were found.

Countries were responsible for testing their national versions based on scripts provided by the Consortium. Scoring testing at the national level was especially important when the correct response included translated and/or adapted textual and numerical information. The testing was done manually, that is, the tester completed each item multiple times, responding to items correctly and incorrectly as specified in the script. That script included the expected scoring result for each response so the tester could compare the observed and expected scoring result. Discrepancies were documented and reported to the Consortium for debugging, with testing iterations continuing until all problems were corrected.

2.3.5 Moving from the Field Test to Main Study instruments

Following analysis of the Field Test data, a number of steps were followed to develop the Main Study instruments.

- **Item analysis**
  Items were evaluated based on their statistical performance in the Field Test, looking at performance within and across countries as well as across modes (i.e., computer and paper). The purposes of the Field Test analyses were to ensure that items were reliable, valid and comparable across countries and that common scales could be developed across countries and assessments.

- **Item selection**
  Based on the Field Test data, developers recommended a draft set of Main Study items for each domain in December 2010. These items were reviewed by the expert groups who, in partnership with developers, finalized the set of items. The recommended set was then presented at a meeting of the NPMs as well as the BPC for their approval.

One challenge for the Main Study selection process was the need to fit the final set of items within the testlets that made up the adaptive design. As shown in Table 1.2 in Chapter 1, the design for the computer-based adaptive instrument included two stages, divided into a total of seven testlets. To accommodate this design, developers needed to look at the difficulty level of items available for the Main Study and determine the appropriate testlets and blocks for the items. For literacy, the fact that items existed as units, or sets of items associated with a single stimulus, posed an additional challenge, particularly in those cases where items within a unit were spread across the defined difficulty levels.

- **Item corrections**
  Countries reviewed the set of items selected for the Main Study, looking for any errors in translation or implementation identified by the Field Test data or during the final national check of those items. Errors were corrected and the final version reviewed and approved for implementation for the Main Study.

The set of items for the Main Study was balanced in terms of construct representation, based on the overall distributions recommendations in the framework. A total of 58 items was selected for literacy and numeracy, with the distribution across linking and new paper and computer versions shown in Table 2.4 below.

Table 2.4: Literacy and numeracy items in the PIAAC Main Study

	Literacy		Numeracy	
	Linking	**New**	**Linking**	**New**
Paper-based	18	6	19	6
Computer-based	30 (including computer versions of the 18 above linking items)	22	28 (including computer versions of 14 of the above linking items)	22 (including computer versions of 3 of the above linking items)

The distribution of these items based on the task characteristics defined in each domain framework is detailed below.

2.3.5.1 Literacy

The distribution of the literacy items included in the Main Study by task characteristics is presented in Tables 2.5-2.7 below.

Table 2.5: Distribution of literacy items by medium

	Final item set		Framework goal
	Number	%	%
Print-based texts	36	62	70-80
Digital texts	22	38	20-35
Total	58	100	100

Note: Each category includes continuous, noncontinuous and combined texts.
.

Table 2.6: Distribution of literacy items by context

	Final item set		Framework goal
	Number	**%**	**%**
Work	10	17	15
Personal	29	50	40
Community	13	23	30
Education	6	10	15
Total	58	100	100

Table 2.7: Distribution of literacy items by task aspects

	Final item set		Framework goal
	Number	**%**	**%**
Access and identify	32	55	40
Integrate and interpret	17	29	45
Evaluate and reflect	9	16	15
Total	58	100	100

2.3.5.2 Numeracy

The distribution of the numeracy items included in the PIAAC survey by task characteristics is presented in Tables 2.8-2.10 below.

Table 2.8: Distribution of numeracy items by response (process)

	Final item set		Framework goal
	Number	**%**	**Number**
Act upon, use	34	61	50
Identify, locate or access	3	5	10
Interpret, evaluate	19	34	40
Total	56	100	100

Note: Each category includes continuous, noncontinuous and combined texts.

Table 2.9: Distribution of numeracy items by context

	Final item set	
	Number	**%**
Everyday life	25	45
Work-related	13	23
Society and community	14	25
Further learning	4	7
Total	56	100

Table 2.10: Distribution of numeracy items by mathematical content

	Final item set		Framework goal
	Number	**%**	**%**
Data and chance	12	21	25
Dimension and shape	16	29	25
Pattern, relationships and change	15	27	20
Quantity and change	13	23	30
Total	56	100	100

2.3.5.3 Problem solving in technology-rich environments

Fourteen PSTRE tasks were included in the Main Study. These included both short and long scenarios.

The distribution of the PSTRE assessment items included in the Main Study by task characteristics is presented in Tables 2.11-2.13 below.

Table 2.11: Distribution of PSTRE tasks by cognitive dimensions

	Number*
Setting goals and monitoring progress	4
Planning	7
Acquiring and evaluating information	8
Using information	6

*Some tasks address more than one cognitive dimension so total is more than 14

Table 2.12: Distribution of PSTRE tasks by technology dimension

	Number*
Web	**7**
Spreadsheet	4
Email	9

*Some tasks involve more than one technology environment so total is more than 14

Table 2.13: Distribution of PSTRE tasks by context

	Number
Personal	8
Work/Occupation	4
Civic	2

2.4 Conclusion

The decision to deliver PIAAC as a computer-based assessment presented both opportunities and challenges for the development of the cognitive instruments. Computer delivery allowed the inclusion of technology-based texts and environments, reflecting the range of materials that many adults encounter in their everyday lives. It also allowed adaptive testing, more reliable computer-based scoring, and the ability to collect a broader range of performance data including timing and process information. One significant challenge was that, in keeping with the open-ended response format used in IALS and ALL, developers needed to define response modes that could allow a reasonable range of open-ended responses while still being computer scored.

The three expert groups considered the implications of computer delivery in their frameworks for literacy, numeracy and PSTRE. Those frameworks defined the general outlines of the assessment instrument in each domain, specifying the task characteristics to be manipulated by test developers and outlining the relative proportion of items to be developed based on the key variables associated with those task characteristics.

Instrument development for the literacy and numeracy domains included selecting linking items from previous large-scale assessments and developing new items. The selection process for linking items involved considering how response modes for items could be adapted to open-ended, computer-scored formats as well as evaluating display and formatting issues for stimulus materials. New items for literacy, numeracy and PSTRE were developed with input from participating countries that included item submissions and a detailed review process. Additionally, developers worked closely with the expert groups who reviewed and, particularly in the case of numeracy, developed items for inclusion in PIAAC. This collaborative endeavor, with input from individuals with a range of expertise and perspectives, resulted in a set of innovative cognitive instruments that provided important information about the skills and knowledge of adults across participating countries.

References

Lennon, M., Kirsch, I., Von Davier, M., Wagner, M., & Yamamoto, K. (2003). *Feasibility study for the PISA ICT literacy assessment.* Princeton, NJ: Educational Testing Service.

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd ed.). Washington, D.C.: American Council on Education.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(1), 13–23.

Chapter 3: The Development of the PIAAC Background Questionnaires

Jim Allen and Rolf van der Velden, ROA; Susanne Helmschrott, Silke Martin, Natascha Massing, Beatrice Rammstedt and Anouk Zabal, GESIS; and Matthias von Davier, ETS

3.1 Introduction

This chapter documents the work done by the Consortium to develop, test and refine the PIAAC BQ. It starts in Section 3.2 by describing the conceptual framework that provides the underpinning for the BQ, outlining the main policy questions that the PIAAC project seeks to answer, and providing the theoretical underpinnings of the concepts that are needed to answer these policy questions and, hence, represented in the BQ. Section 3.3 briefly explains the rationale underpinning the JRA module in the BQ, which was developed separately from the main master BQ by a different team of experts. Section 3.4 deals with the development and validation of the BQ, including an outline of the decision-making process for selection of items in the BQ, a brief summary of two rounds of cognitive testing that were conducted, and a report on the analysis that was conducted of the data from the Field Test with a view to refining and shortening the BQ for the Main Study. In Section 3.5, a brief outline will be given of the content of the BQ, including an overview of the structure, and a brief description of the national adaptations and extensions that were made. Finally, in Section 3.6 we explain how the BQ was implemented on the TAO platform, in particular in terms of the use of instructions to interviewers, help buttons and consistency checks that allowed the BQ to be administered in a coherent and standardized way across the participating countries.

3.2 The PIAAC conceptual framework for the BQ

The policy questions

The PIAAC project seeks to answer the following policy questions:

A. How are skills distributed?
B. Why are skills important?
C. What factors are related to skill acquisition and decline?

How are skills distributed?

Human capital is considered the driving force of economic growth. Investments in skills are vital to keep up with technological change (the so-called Skill-Biased Technological Change) as well as other changes resulting from market or organizational developments (e.g., the introduction of High Performance Workplace Practices). Policymakers have an interest in monitoring the stock of human capital in their country and identifying the different levels among relevant subgroups. PIAAC assesses the stock of human capital in a society by providing a descriptive analysis of the distribution of skills proficiencies and skills use in the adult population. The survey enables countries to answer questions such as:

- How does the adult population in a country compare to that of other countries in terms of average levels of skills proficiency and skill use? What share of the adult population has low proficiencies of relevant skills?

- In terms of equity, how are skills distributed among relevant subgroups, such as gender, age group, region or migration status? Are certain subgroups particularly vulnerable to low skills proficiencies?

- How are skills proficiencies distributed across sectors of industry? Are there certain sectors of industry that are characterized by particularly low levels of skills proficiency? How do the skill levels of these sectors compare to those in other countries?

- How are skills proficiencies distributed across different levels of schooling when benchmarked against other countries? Are there population subgroups that appear to be underserved by the current education system? Is there an underdevelopment of skills at particular levels of education? What are the skill levels of early leavers from education?

- Who is participating in adult learning of various types? To what extent are particular population subgroups excluded from adult learning systems?

Why are skills important?

There is little interest from a policy point of view for any investment in skills if it has no relation to relevant outcomes. Other services are competing with education and training for a share of budgets, so the case for returns to educational investment needs to be made on a secure and sophisticated evidence base. Moreover governments and the public make education accountable to show the effects of their efforts. For that reason, one of the key goals of the BQ is to provide indicators that can be used to show if differences in skill matter economically and socially. The most obvious area in which policymakers are interested is how skill levels are related to economic outcomes of individuals. Cognitive skills are thought to be a key determinant of an individual's productivity, and therefore it is not surprising that cognitive skills are related to economic success. There is a large body of evidence showing that higher cognitive skills are indeed associated with better labor market outcomes. Relevant questions are:

- How are skills related to individual employment opportunities and job security?

- How are skills related to earnings and other indicators of labor market success?

- Do low skill proficiencies form a barrier to individuals entering the labor force?

- Are low-skilled people more affected by job insecurity? Is there a minimum level of skills needed to be employable?

- How do skills affect the relation between education and training on the one hand and economic outcomes on the other? Can skills compensate for low educational qualifications?

Apart from economic outcomes, other areas are of interest as well, such as the relation among skills, health status and civic participation. Adverse outcomes in such areas place large burdens on governments, businesses, and individuals, including both the direct expenditure of resources (such as government spending on health care) and indirect costs (such as the value of goods and services workers do not produce while ill).

Relevant questions are:

- To what extent is literacy related to health status of individuals, various subgroups, and the overall population?

- To what extent do individuals with low skills appear to be less engaged in the broader society (voluntary work, social trust)?

- How do individuals with low skill levels cope with their everyday reading and numeracy demands? To what extent do these coping mechanisms make these individuals reliant on others? To what extent does the engagement of migration groups or linguistic minorities appear to be inhibited by their lack of skill in the language of the test?

- Do high-skilled people have a higher involvement in civic activities? What is the relation between skills and the level of social trust?

What factors are related to skill acquisition and decline?

Under the assumption that skills matter economically and socially, policymakers have an interest in knowing what factors are related to higher skill levels. Of course, the prime focus is to assess the effects of factors directly affected by policy, such as the provision of formal and non-formal organized learning activities like education and training. But it is also relevant to compare the efficiency of these skill production routes with the efficiency of others not directly under the control of policymakers, such as the informal learning activities in which people can engage. Assessing the overall relation among education and training and skill levels is only a first step in unraveling the determinants of skills acquisition. We can assume that not all education and training activities have the same impact on skills development. Nor can we assume that the impact is the same for all relevant subgroups. Policymakers have an interest in seeing which characteristics of education and training are most strongly related to higher skill levels in the population and which subgroups appear to profit most from which type of intervention. Finally, we need to be aware that skills can be acquired, but also can be lost. Preventing skill decline is probably just as important as promoting skill acquisition, but the underlying factors affecting these processes may be quite different, and it is important to have good insight in both processes.

For these reasons, the survey was designed to enable countries to answer questions like:

- What is the relation between education and training and the skill development of people? Are these relations different from those with other learning activities that people engage in to develop their skills, such as informal on-the-job learning?

- Are the effects of education and training the same for each subgroup? Are there subgroups that appear to profit from the investment in education and training?

- What is the relation between underinvestment in work-related training and adult skill levels? How are characteristics of the work environment related to skill levels? Is informal learning an on-the-job a substitute for work-related training?

- How do processes of skill acquisition and decline vary with age? What are the factors related to skill decline? Are these the same factors as are related to skill acquisition?

Theoretical background

In this section we describe the main theoretical elements of the conceptual framework and, where relevant, indicate the items that have been included in the BQ to reflect these elements. The purpose of this part was to provide a solid theoretical basis for the policy questions formulated in the previous section. It also served as a guideline for the selection of relevant concepts and the translation of those concepts into specific questions in the BQ. This framework also served as a guideline for the analysis and interpretation of the data in the Field Test, where it was used to derive predictions on how particular sets of variables were expected to behave. Its main function in the Main Study is as a basis for deriving hypotheses pertaining to the policy questions outlined in the previous section.

The presentation of the theoretical framework will be divided into three parts, roughly corresponding to the three types of policy questions described above. We start with a brief overview of the literature on the nature of key skills. Although the direct assessment (DA) as such falls outside the scope of the development of the BQ, the raison d'être of the BQ is to provide the context information needed for analyzing and interpreting the results of the DA. As a consequence, it is essential to proceed with a solid understanding of what is being measured in the DA and, equally important, what is not being measured. We then summarize the literature pertaining skills acquisition and decline. The theoretical discussion is concluded with a review of the literature on outcomes of skills.

What are key skills?

As noted above, policymakers have a strong interest in knowing how skills are distributed across countries as well as across different subgroups within countries, such as age, gender, ethnicity, regions, sector of industry, and levels and fields of education. If we want to answer these questions, it is important to first take a step back and reflect on what is being compared. Below is a brief overview of the literature on so-called key skills, of which the skills measured in PIAAC form an important subset.

The quest for key skills

The last few decades have seen an increased awareness of human capital as one of the driving forces of economic development. Policymakers have realized the importance of investing in education and training as a way of improving the existing stock of skills. This has resulted in an accompanying need to monitor and assess the stock of human capital. What soon became clear is that education as such is a poor indicator of the stock of human capital. Individuals with the same nominal level and type of education can differ markedly in their command of various skills. Likewise countries that have more or less comparable levels of educational attainment can nevertheless differ substantially in the level of skills that are acquired in education. This has been shown in studies like ALL and the Programme for International Student Assessment (PISA).

As the emphasis shifts from educational qualifications towards skill measurement, the question naturally arises as to what skills should be measured. It seems clear that in order to perform even the most basic tasks, many discrete skills are required. Determining which skills should be measured is a complex and difficult task, which is compounded by the fact that people not only make use of generic skills such as the ability to communicate or the ability to learn, but also of a large number of highly specific skills pertaining to particular tasks, situations and objects.

In order to introduce some order in the understanding of the diversity of human skills, many scholars have engaged in a quest for so-called core skills or key competencies. A major project in this respect was the DeSeCo (Definition and Selection of Competencies) project. This project was initiated by the OECD to provide an overarching framework for international skills assessments. Competencies are defined in this project as "the ability to successfully meet complex demands in a particular context through the mobilization of psychosocial prerequisites (including both cognitive and noncognitive aspects)" (Rychen & Salganik, 2003, p. 43). The basic difference between this view and earlier concepts of skills is the holistic nature of the concept of competence. It refers not only to a range of cognitive and noncognitive skills and other prerequisites that need to be in place in order to perform in a competent way, but also to the notion of "orchestration," which is defined as the ability to use these constituent elements in a meaningful and deliberately arranged way.

Although the theoretical framework provided by the DeSeCo project injects some welcomed theoretical rigor into the discussion of skills measurement, it does not in itself directly give rise to clear recommendations as to the competencies to be measured. The best way to conceive of this overarching framework is to see that it indicates the main underlying competencies that give skills their significance.

Binkley et al. (2003) developed a framework that provides more detailed guidance for the development of skills measurements. This work concentrated on two strands of research: what skills are necessary in the workplace, and cognitive functioning. From the first strand, a list of six skill areas was extracted that seemed to underlie many of the most important skills: communication (speaking, listening, reading and writing), mathematical, problem solving, intrapersonal (motivation, metacognition), interpersonal (teamwork, leadership) and technology. From the strand of psychological theory, four core domains of intelligence were extracted: practical abilities, crystallized analytical abilities, fluid analytical abilities, and creative abilities (the ability to cope with novelty). As the authors point out, the two strands are not mutually exclusive, but rather represent different aspects of skill. The workplace skills provide the context within which each of the four core intelligence domains are expressed; conversely, each category of workplace skill can involve four distinct types of thinking.

The choice of direct assessments in ALL was based not only on these theoretical notions but on practical considerations such as an established tradition of measurement where assessments are sufficiently compact to be used in a household survey. As a consequence, ALL concentrated on only part of the matrix formed by the intersection of the two strands of research, in particular the more generic aspects of the communication and mathematical skill areas. PIAAC builds on the direct assessments in ALL, extending these to the area of problem solving in technology-rich environments, which contains elements of the problem solving and technology skill areas. Although it is not possible to draw any sharp dividing line, the three domains of direct assessments in PIAAC differ in the extent to which they relate to the four types of thinking derived from psychological theory. Because the developmental pattern throughout life is thought to be quite different for the different types of thinking, this has important implications for the manner in which the different skills can typically be expected to be acquired and in some cases eventually lost. We will return to this point below.

To the extent that the skills measured in the direct assessments are shown to be related to important economic and social outcomes (see below), the pragmatic restriction to those skill aspects that lend themselves well to a survey approach need not seriously diminish the value of the information gathered. It is important, however, to keep in mind that we are dealing with a subset of the skills possessed by the individuals participating in the survey. The intrapersonal and interpersonal skill areas are not included in the direct assessment, but as will be outlined below, these are covered to some extent by items included in the BQ. Arguably the most conspicuous omission is in the area of specific skills used by individuals in their chosen line of work.

The importance of professional expertise

Even though employers often list generic cognitive skills and personal traits skills as the most important ones required in the workplace, professional expertise is a condition sine qua non for success in many occupations. For example, nobody would doubt that in order to become a good medical doctor, architect or car mechanic, one needs to acquire the domain-specific knowledge and skills that make up the professional domains of these occupations. The German psychologist Weinert formulated this as follows: “Over the last decades, the cognitive sciences have convincingly demonstrated that context-specific skills and knowledge play a crucial role in solving difficult tasks. But generally, key competencies cannot adequately compensate for a lack of content-specific competencies” (Weinert, 2001, p. 53).

There is, however, a plethora of specific professional skills. It is not be possible to measure professional expertise directly in the PIAAC assessment, simply because there is no common assessment instrument that allows all different types of professional skills to be measured in a meaningful way for large populations. The absence of direct measures of specific skills underscores the importance of obtaining information on the occupation of working respondents, based on the answers to questions D1a and D1b in the BQ. As the differences among occupations in the skills measured in the direct assessments is likely to be at least matched and probably eclipsed by differences in level and type of specific skills, the residual occupation-level variance in economic outcomes should provide a rough indication of the economic importance of specific skills relative to the generic skills measured.

Although no direct assessment of occupation-specific skills is included in the PIAAC survey, measures of skill use in some more generic work-related areas, as well as in the domains covered by the direct assessment and in the area of interpersonal skills, have been developed in a separate module based on the JRA. This module is described in Section 3.3.

Current investments in education and training

From a descriptive point of view, it is important that PIAAC provides accurate information on current levels of education and training. Access to lifelong learning by different groups remains a crucial issue for governments of the OECD member countries. Formal education (B_Q01-B_Q10), formal training (B_Q12-B_Q20), and informal training (D_Q13a-c) all contribute to the stock of human capital, and countries will display different profiles in how the human capital stock is built up. PIAAC will provide a snapshot of human capital investments by the incidence and intensity of training during the previous 12-month period. From a policy viewpoint it is important to not only obtain an indication of the volume of investments, but in the case of adult education and training, to have information on financing of such investments. A large part of adult education and training efforts are paid for by employers. Since most training received by individuals also benefits other employers (externalities of training) this typically leads to too little work-related training being provided because part of the returns are captured by outside parties (competing organizations and the individual). From a policy perspective, this could warrant some interventions in the training market to balance out a potential source of underinvestment in training. In addition, knowledge on current investments in learning can contribute to the formation of policies designed to provide more equitable or effective inducements to encourage participation among those most in need of further learning. This refers both to differences across different skill levels (Are low-skilled individuals investing enough in their human capital?) and across key reporting categories as specified below. The questionnaire contains indicators of whether the training was followed in working hours (B_Q15b, to assess the level of investment by employers in training in terms of opportunity costs), whether the respondent's employer contributed to the costs of training (B_Q16, to assess the level of direct investment in training by employees, employers and other actors), and (reasons for) nonparticipation in learning activities in which the respondent would have preferred to engage (B_Q26a-b).

When analyzing training, it is necessary to be able to distinguish different categories of training. At the most general level, it is important to distinguish work-related from non-work-related training (B_Q14a). Work-related training is usually expected to have some effect on performance, which is presumably expected to be based on increased skill levels, and to result in productivity and possibly wage gains. Training that has been undertaken for other reasons may also increase certain skills but would not necessarily lead to productivity increases at work.

Reporting categories

For reasons of effectively addressing skill deficiencies, but also from the point of view of social equity, it is important to have a good picture of where the deficiencies are most concentrated. Are there population subgroups that appear to be underskilled? To answer these questions, we need to know how skills are distributed among relevant subgroups, as defined, for example, by gender (A_N01a), age (A_Q01a), socioeconomic background (J_Q06b, J_Q07b) or migration status (J_Q04, J_Q06a, J_Q07a). These so-called reporting categories are important both from a point of view of equity and efficiency: If skill gaps lead to social and/or economic exclusion, this is not only detrimental to the well-being of the groups involved, but also to the functioning of the economy and society. Because the reasons for skill gaps are likely to be systematically different for different “at risk” groups, the policy measures undertaken are likely to be group-specific. Age is also important because both skills acquisition and skills decline are related to age, leading to typical age profiles of skills and skill-related outcomes.

Region (collected through the Case Management system) is an important reporting category as well because of strong regional differences in level of economic development in some countries. It may be that certain regions are being held back by particularly low levels of skills proficiency, or conversely, that regions can be identified where skill demand is particularly low. In addition, because policy is often formulated and/or implemented at the regional level, it is crucial to have access to outcomes at that level.

Occupation (D_Q01, E_Q01), sector of industry (D_Q02, E_Q02) and firm size (D_Q06, E_Q06) are needed to detect areas in which skill gaps exist and to assess the extent to which training investments are taking place to reduce these gaps. This and similar information form the basis for directing possible policy interventions to those groups where intervention is most needed.

Because highest level of education (B_Q01) is assumed to be one of the strongest predictors of skills (see below), and because this is differentially distributed across countries, a breakdown by this variable is needed for even the most elementary understanding of the results. In addition it is important to know how access to the education system is distributed across different subgroups that are “at risk” from the point of view of skills proficiencies.

Determinants of skills acquisition and decline

As was the case for defining and measuring skills themselves, there is not just one but several strands of research pertaining to how individuals acquire and in some cases lose skills over their lifetime. One prominent strand is that of the economics of education. Since the pioneering work by scholars such as Becker (1964) and Schultz (1963), economists have looked at education, training and other activities undertaken by individuals to improve their level of knowledge and skills as investments in human capital that are expected to yield returns in the labor market. A second major strand is that of sociological research that points to the social environment affecting school choice and educational attainment. The third strand is educational research, in which scholars have tried to uncover those features of education that are particularly effective in promoting learning. Fourth, a conceptually related but empirically largely distinct area concentrates on how people continue to learn after leaving initial education. An important focus of this strand of research is on courses, workshops and other forms of training in which employees participate, but in recent years the focus has increasingly broadened to include features of the job or organization that promote informal learning. Finally, this focus on lifelong learning has led to increased attention to the fact that individuals not only acquire skills over their lifetime but are also confronted with skill loss and a general decline in the ability to acquire and retain new knowledge and skills. In this section we will look at each of these strands of research in turn.

Education as an investment

In economics, education and learning are treated as an investment. From this point of view, people are expected to invest in education and learning when the costs are smaller than the future benefits. Not everybody is equally likely to invest in the same amount of education. People differ in the degree in which they enjoy education or learning and in the degree to which they value the potential benefits of education. Due to heterogeneity in preferences, there will also be heterogeneity in the decision to learn. Borghans et al. (2007) provide a model for investments in education and learning that capture a wide range of potential differences between individuals.

First, people differ in their capacity to acquire skills. The costs of education are lower for people who acquire skills more easily because they learn faster. The capacity to learn depends not only on innate cognitive abilities but also on personal traits. For example, someone who is easily distracted from a task will need more time to learn. Second, people differ in preferences. They might differ in how they value learning, working and leisure. They might differ in how much they value a high income or other potential benefits of education, and they might differ in how they value future benefits compared to current benefits (time preference, the discount rate) and how they account for risks in outcomes (risk aversion). Third, people might face constraints in their choices. Credit constraints can influence the decision to attend school, but also a lack of facilities for education and less favorable family conditions can be treated as such constraints. Finally, the decision to invest in education will depend on information available at the time of investment. If people don't know about the benefits of education, it is unlikely they will invest.

The main reason it is important to take account of factors expected to influence willingness to invest in education is that they may have a direct impact on skill levels distinct from the indirect effect via the increased level of investment in education. If such factors are not taken into account, estimates of the effect of education on skill levels will be biased. The BQ covers some, but not all, of these factors. The questionnaire contains no direct indicators of innate learning abilities. It does, however, include a number of control variables that are related to this concept, in particular the family background in terms of parents' education (J_Q06b and J_Q07b). Learning strategies (I_Q04) are included as they may affect individuals' ability to learn.

The social environment

The constraints facing different social groups have been extensively studied by sociologists, who have a long tradition of research looking at the social barriers to education and training. While gender inequality in initial education has vanished and actually turned into an advantage for girls in many Western countries, it still persists in occupational careers and later access to training. The sex of the respondent is therefore a key reporting category for PIAAC. Inequality in access to education related to the family background both in terms of socioeconomic status and migration status is more persistent.

Part of these differences relates to differences in school performance and learning abilities, the so-called primary effects of social stratification (Boudon, 1974). These may be caused both by differences in innate abilities and socialization processes. The cultural capital of the family (Bourdieu, 1984) in particular provides a powerful predictor of the school performance. But even with the same school performance, students from different family backgrounds make systematically different choices in education (the secondary effects of social stratification), and given the number of choices that have to be made during the educational career, the cumulative effect of these choices might even overwhelm the primary effects. These differences in choices relate to differences in social cost-benefit analyses. The social costs and benefits involved in obtaining education are different for students from different social backgrounds. Following an educational career that is different from the one that is common in the family induces social costs, while the social benefits may be lower. The BQ includes indicators of gender (A_N01), parents' education (J_Q06b, J_Q07b), migration status (J_Q04a-c, J_Q06a, J_Q07a), cultural capital in parental home (J_Q08), and language used in parental and current home (J_Q05a1-2, J_Q05b).

Effective learning and instruction

Following a certain type of education or training path does not automatically imply that all students are likely to acquire the same set of skills. Educational research has shown that there is considerable variation among educational systems, schools, study programs and teachers in how much skills students acquire during education or training. A large part of the effect of education on skill development is likely to be indirect, as students are turned into more effective or less effective learners for life. In other words, different characteristics of education may affect both the direct acquisition of skills as measured in the direct assessments, as well as the ability to acquire these skills after leaving education. Without providing too much detail, we can note a number of interesting approaches here:

- Situated learning theories (Glaser, 1991) emphasize that competencies and competence development are context-specific. They stress the importance of coherence and context-relevance (e.g., real-life experiments, simulation and practical work experience) in the design of the curricula in order to develop expertise.

- Active learning theories reject the traditional naïve model of the teacher as the expert, imparting his or her knowledge directly to the student. "Powerful learning environments" (De Corte, 1990) and active instructional methods like problem-based learning and project-oriented education are thought to foster the development of generic competencies like problem solving and metacognitive abilities.

- In addition to these innovative ways of learning based on elaborate theories on how individuals actually learn, educational research has traditionally stressed "time on task" as one of the most important factors affecting student outcomes. That is, the actual time students spent on education (within the classroom and through self-study) is a good predictor of the learning outcomes net of other factors.

Although it is not practicable to describe the educational environments respondents have been exposed to, it does make sense to include indicators of respondents' learning strategies, which may in part be a result of such exposure. As Peschar (2003) has remarked, such strategies can be seen as important prerequisites for learning throughout one's life. Self-regulated learning theories point to the relevance of metacognitive abilities and information-processing strategies of students (Kolb, 1984). Learning styles differ among students, ranging from memorizing and rather atomistic ways of learning towards a more constructivist approach in which concepts and theories are actively incorporated in a coherent body of knowledge. Although such attitudes are likely to be heavily influenced by one's family background, either directly through genes or indirectly through early socialization, there is evidence that such attitudes and strategies can be influenced by education. Question I_Q04a-m contains indicators of learning strategies. Although the list of items has been strongly based on previous international comparative research, the question in its current form is new.

Among the characteristics of the educational career, the achieved level of education (B_Q01a) is, of course, the most important concept affecting skill levels. More years of schooling are expected to have a positive impact on the skills proficiencies. Based on the information of national experts, all reported national categories in the achieved level of education are converted into the nominal years of schooling needed to achieve that particular level of education (see Appendix 5). Moreover, the particular field of education (B_Q01b) followed will also affect skill levels: Graduates from certain fields of education will have higher scores in the literacy domain; others will probably have higher scores in the numeracy domain.

Other relevant characteristics of the educational career that may affect the skills development are the type of pathway in secondary education (whether a general or school-based vocational (B_Q01a). Based on the information of national experts, we determined for all relevant reported national categories in International Standard Classification of Education (ISCED) Levels 2 to 4 whether the types of pathway in secondary education was general or vocational (see Appendix 5). It is also important to identify whether the education has been completed outside the host country (in the case of migrants) in order to identify any negative effect on literacy skills. The BQ therefore contains information on where the highest qualification was obtained (B_Q01a2).

Training and informal learning

People do not only learn during initial education but later in life. In the human capital literature, many studies have analyzed the effects of workplace training participation on workers' wages (see Bassanini, Booth, Brunello, De Paola, & Leuven, (2005) for an overview). Several studies have found high returns on workers' participation in training. Brunello (2004) found that having recently attended training increases a worker's income by about 12 percent.

However, one may wonder whether it is really the participation in formal training that makes the difference. Borghans, Golsteyn and de Grip (2006) show that employees spend much more time on informal learning activities than on formal learning. They also found that when employers stimulate workers' participation in formal courses, these workers will also spend more time on informal learning in the workplace. As many of the studies on the effect of formal training do not measure the time spent on informal learning, all the benefits of the knowledge and skill acquisition of the workers are attributed to their participation in formal training. It is important that PIAAC not only looks at the incidence of formal training but also explores various kinds of informal learning, as they contribute highly to skills acquisition.

Arrow (1962) emphasized the importance of unstructured workplace learning, not from the perspective of the individual worker but that of the firm. He found that informal learning is a more or less automatic byproduct of the regular production process of a firm, which he labeled "learning by doing." Furthermore, job characteristics might also affect post-initial schooling. Employees with mainly monotonous tasks are expected to attend less formal training than those in jobs with more complex tasks. Jobs that require problem solving and learning new things probably include high training incidence and informal learning as well.

Human resources practices and job characteristics are the major work characteristics that determine the opportunities for workers to attend training and learn in an informal way. Although these opportunities are often necessary for actual training behavior, a workplace characterized by these training opportunities might not be sufficient. Workers' characteristics will probably determine whether the learning opportunities at work are fully exploited. Personal characteristics such as age, gender and level of schooling are found to be important determinants of post-initial schooling behavior (Bassanini et al., 2005).

In addition to measures of participation in education at the time of the survey and over one's lifetime (B_Q01 to B_Q10), the BQ contains questions on recent investments in training (B_Q12 to B_Q20), including the main reason for participating in training (B_Q14b), crucial for analyzing the effects of training, informal training by supervisors, colleagues, etc. (D_Q13a), learning by doing (D_Q13b), keeping up to date with new products or services (D_Q13c) and work autonomy (D_Q11a-d).

Regardless of the specifics of the training and learning practices applied in the organization in which individuals work, the amount of work experience acquired can be expected to have a strong effect on skills development. In wage estimations, work experience is generally assumed to be positively related to productivity, but the returns are assumed to diminish with further experience. In terms of skill acquisition, this is consistent with the notion that one is likely to be most exposed to situations from which one can learn something new early in one's career. As the career develops, the chance that one will be exposed to new stimuli is decreased. This pattern is likely to be reinforced by typical patterns of brain development over the lifecycle, which predicts a steady decline in learning and retention abilities from young adulthood onward. We will return to this point below.

Skill acquisition is not only dependent on total experience, but also on the specific way in which this experience has been acquired. In addition to total work experience, the number and timing of changes of employer and/or career breaks is therefore also important. There is probably a certain minimum time one would need to remain with a given employer to have a reasonable chance of learning new things, and the returns to tenure in most jobs are likely to remain positive for at least a few years (although probably not in very low-level routine jobs, see below). Because the new experiences one can expect to be exposed to when working for a given employer are likely to diminish over time, we would expect a certain number of job changes over the career to have a positive effect on learning. Lengthy career breaks comprise periods during which the exposure to work-relevant experiences is likely to be limited.

In addition to these direct effects of work experience on learning, there may be indirect effects when work history is interpreted by potential employers as a signal of productivity and learning potential. In that case, a career characterized by frequent changes and/or lengthy interruptions may affect the willingness of potential employers to hire an individual and to invest in his or her human capital. Lengthy periods of unemployment – that is, seeking work without success – may additionally exert a negative effect on individual motivation.

The questionnaire contains a number of questions related to the above-mentioned aspects of employment history. Question C_Q09 allows us to establish the total number of years of work experience (if any) the respondent has acquired in his or her lifetime. Question C_Q10a provides information on the number of different employers worked for in the last five years.

Skill loss

The increased focus on lifelong learning in recent years has led to increased attention to the fact that individuals not only acquire skills over their lifetime but are also confronted with skill loss and a general decline in the ability to acquire and retain new knowledge and skills. The single-most important finding of IALS and ALL was that skill loss was large enough to offset all of the expected gains from increasing educational quality and quantity. Until now, only scattered studies on different aspects of skills obsolescence have been published. Most of these studies were published in periods in which unemployment was high. This increased the focus on the adverse impact of skills obsolescence for the workers involved. It is interesting that in the recent policy debates on skills obsolescence and "lifelong learning," the main focus has been on the waste of valuable human resources and on the nonoptimal performance of workers with inadequate skills. This brings skills obsolescence to the heart of the economic challenge the western economies face: in realizing the transformation towards a knowledge-based society with an aging population.

From a cognitive and neuropsychological perspective, higher order brain functions follow a steep developmental pattern and reach a plateau of optimal functioning in young adulthood. Such processes and changes therein can be measured on a behavioral level using dedicated neurocognitive instruments which tap the efficiency within specific neuropsychological domains, such as language, intelligence, memory, attention and speed of information processing.

Optimal neurocognitive development is dependent on a complex interplay of factors, with genetics, socioeconomic status, educational achievement, adequate nutrition, and uncompromised mental and physical health being the strongest predictors of developmental success. Researchers have coined the term "brain reserve capacity" (BRC) to indicate the neurobiological constraints which determine maximum processing capacity of higher order brain functions. This concept has proven its validity in, for example, predicting individual cognitive aging trajectories later in life. Important proxy measures of BRC include educational level and occupational achievement.

On a population level, most cognitive abilities such as memory function, information processing speed and attention capacity tend to decline with advancing age. Adequate preservation of cognitive abilities is of primary importance to older people, as cognitive decline can result in a loss of productivity among those still working, and a loss of independence and autonomy for retired people. Large individual differences exist in the offset and rate of decline of specific cognitive functions. We drew attention above to the theoretical distinction drawn in psychological research between "fluid" and "crystallized" abilities. The former refers to functions that involve controlled and effortful processing of novel information (cognitive mechanics), and the latter to the representation of learned skills and access to knowledge (cognitive pragmatics). Fluid abilities are far more sensitive to aging (Figure 3.1), and both cognitive domains show different developmental patterns across the life span. Fluid abilities typically start declining in the mid-20s, while crystallized skills may improve until and beyond even the age of 70. The distinction between the two is important because the direct assessments in PIAAC will differ in the extent to which they relate to crystallized or fluid abilities. One may hypothesize that numeracy and literacy skills relate more strongly to crystallized abilities, while dynamic problem solving in a technology-rich environment will relate more to fluid abilities. For adults, the decline in fluid abilities is more likely to strongly hamper their working and everyday life than the decline in crystallized abilities.

Figure 3.1: Theoretical representation of 'crystallized' and 'fluid' abilities over the life span

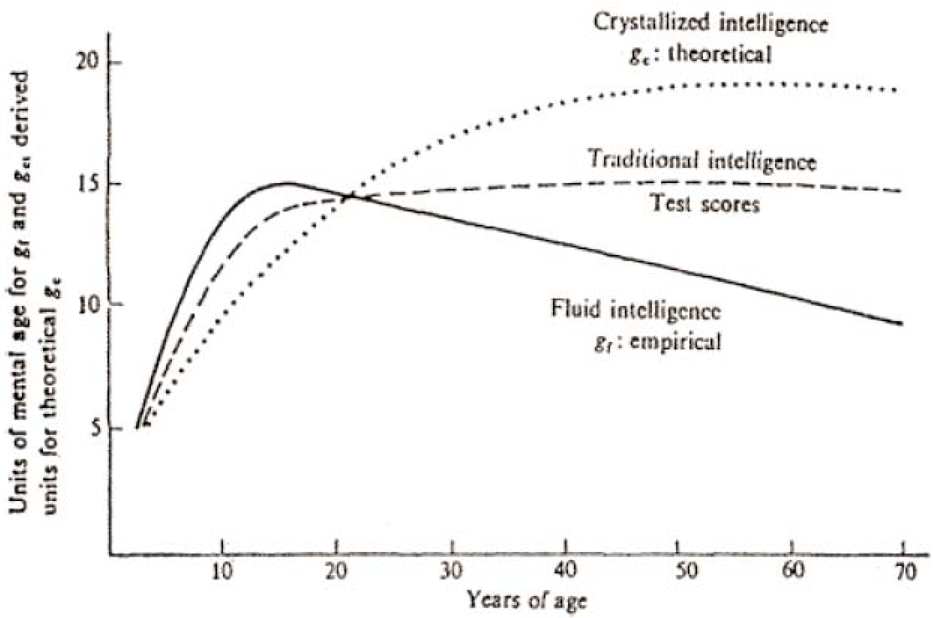


Figure 1. A theoretical description of life span curves of intellectual abilities. From *Intelligence: Its structure, growth and action* (p. 206) by R. B. Cattell, 1987, Amsterdam: North-Holland. Copyright 1987 by Elsevier Science Publishers. Reprinted with permission.

The two most prominent symptoms of "usual" cognitive aging in daily life are a gradual reduction in memory retrieval and information processing speed. Stored information remains relatively intact, but access and retrieval becomes increasingly difficult for older individuals. Another feature that has received considerable interest in research is the reduced ability of older individuals to suppress or inhibit irrelevant information, making decision processes more complicated, and therefore slower.

Still, cognitive aging is not merely a predestined process which ultimately leads to pathological states, such as a cognitive disorder like dementia. The ability to learn new skills is still present in older individuals, but – on average – more time is needed to develop the same level of mastery as for younger persons. Recent advances in cognitive neuroscience have convincingly demonstrated that healthy brains show considerable capacity to compensate for reduced integrity of functional networks or to reorganize existing networks to adapt to changing task demands. The importance of adequate and continued exposure to environmental stimuli during the lifetime is now considered pivotal for optimal conservation of cognitive abilities in old age (conceptualized in the "use it or lose it" paradigm).

Empirical findings suggest that complex intellectual activity increases cognition of older workers (Schooler et al., 1999). Skill investments made during working life might improve people's capacity to continue learning and adapting to new environments. Other factors that are conjectured to affect the development of cognitive ability at later stages in life include occupation, leisure activities, lifestyle and social interaction.

Building partly on such insights from cognitive and neuropsychology, De Grip and Van Loo (2002) developed a typology of different types of skills obsolescence. First, the depreciation of human capital may simply be caused by the wear of skills, resulting from the natural aging process. Physically or mentally challenging working conditions may accelerate the wear of a worker's skills. Large epidemiological studies have shown that health-related factors are involved in the enhanced cognitive decline seen with increasing age. In addition, several chronic diseases have been associated with a reduced cognitive capacity in both epidemiological surveys and clinical case-controlled studies.

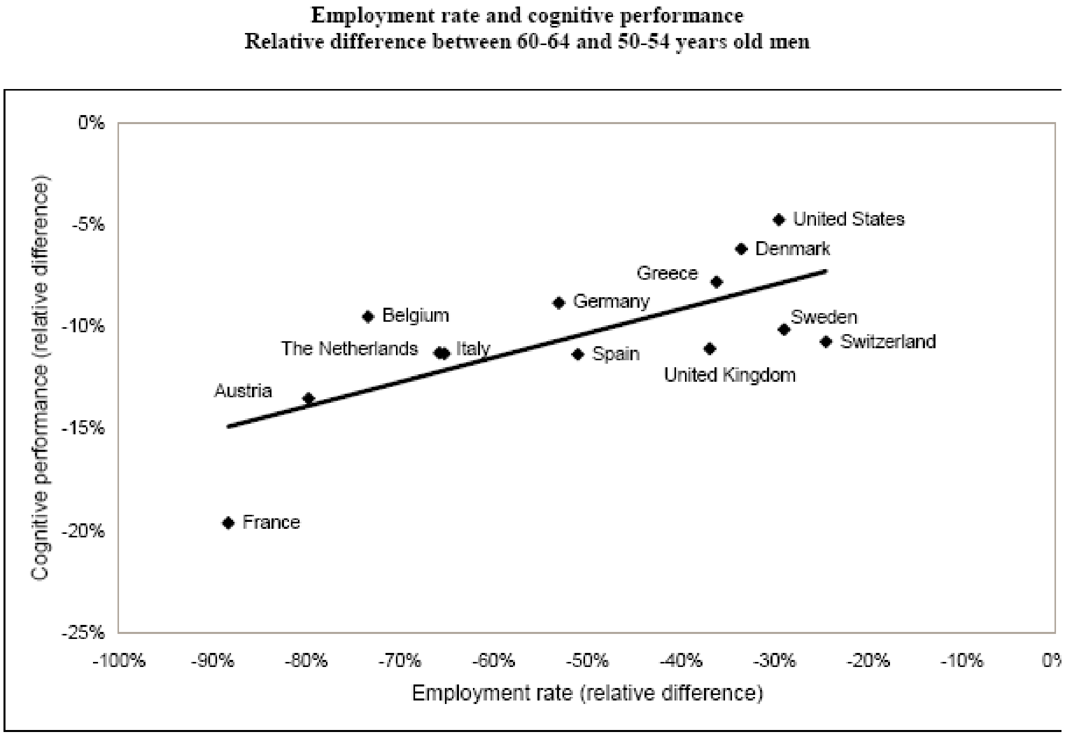
The second category of technical skills obsolescence concerns the atrophy of skills due to the lack or insufficient use of them. This atrophy could result from unemployment and career interruptions, or from employees working below their attained level of education. Arthur et al. (1998) conclude, on the basis of a meta-analysis from the psychological literature on skill decay and retention, that there is substantial skills obsolescence when they are not practiced or used. De Grip et al. (2008) show that job-worker mismatches induce cognitive decline with respect to immediate and delayed recall abilities, cognitive flexibility and verbal fluency. Also, as a result of specialization, certain knowledge and skills acquired during initial education may get lost. Apart from these two factors related to the personal characteristics of the worker, skills obsolescence may also occur as a result of changes in the demand for skills, due to, e.g., technological or organizational developments in the production process.

The BQ enables insight as to some of the possible causes of skills obsolescence, such as age (A_Q01a-b), health (I_Q08), unemployment (C_D05), working below one's level (D_Q12a-b), long tenure (D_Q5a1-2) and sector of industry (D_Q02a-b).

Institutional factors

There is a need to study whether policy and institutions can affect the process of cognitive decline. It is well established that early retirement decisions are largely driven by institutions. Gruber and Wise (2004), for example, show there is a very strong cross-country relationship between retirement rates and government policy. If keeping workers active can postpone cognitive decline, there is an important role for policies that increase labor market participation of older workers. Using data from the Survey on Health, Ageing and Retirement on cognitive skills of the population aged 50 and over, Adam et al. (2006) show that relative average cognitive skills among older workers are on average higher in countries in which – as a consequence of national institutions – participation rates of older workers are also higher (see Figure 3.2).

Figure 3.2: Employment rate and cognitive performance



Source: S. Adam, E. Bonsang, S. Germain and S. Perelman (2007), "Retirement and cognitive reserve: A stochastic frontier approach applied to survey data," CREPP DP 2007/04, University of Liège.

Even though it is extremely important to better understand how the process of cognitive decline can be stopped and whether there is scope and need for policy intervention, the study of the determinants of cognitive decline is still in its infancy. Much can be learned from relating differences across countries to cross-country differences in policies, regulations and institutions. PIAAC offers a unique opportunity to gain such insights as it provides detailed data of the distribution of skills across age. By linking this type of data to information from other data sources on institutional factors, we can at least explore how these relations look at the aggregate level of countries.

Skills and outcomes

We remarked above that the policy relevance of measuring skills is strongly dependent on their effect on relevant outcomes. In addition to economic outcomes such as employment opportunities and rewards in the labor market, it is important to take account of outcomes in other areas that may also be influenced by skills, such as health status, voluntary work, and social trust.

Skills and labor market outcomes

Cognitive skills are a key determinant of an individual's productivity, and therefore it is not surprising that cognitive skills are related to economic success. There is a large body of evidence showing that higher cognitive skills are associated with better labor market outcomes (e.g., Heckman et al., 2006). The most basic of economic outcomes is an individual's current labor status, which is constructed using several questions in the questionnaire (C_D05). A distinction can be drawn between those who participate in the labor force and those who do not. The former category can be divided in turn into those who are employed and those who are unemployed (that is currently not working but available for and actively seeking work). There are several reasons why an individual might fall into the latter category – for example, study, household duties, or sickness or disability. To provide a broader indication of respondents' current situation, in question C_Q07 respondents are asked to report their own self-declared main labor status.

For those currently or recently in work, several important labor market outcomes are included in the questionnaire, including working hours (D_Q10), individual earnings (D_Q16-18), job security (D_Q09), occupational status (D_Q01a-b), and the quality of the match between education and work (D_Q12a-c)

One of the interesting questions in this respect regards the precise role of education and skills in producing these outcomes. There are rivaling hypotheses on this point. Very often the strong relation between education and labor market outcomes is explained in terms of human capital theory (Becker, 1964), which claims that people with more years of schooling earn more because the competencies they acquired in education have made them more productive. While this is probably true to some extent, at least in the aggregate, it tells only part of the story. Scholars such as Spence (1973) and Arrow (1973) have pointed out that the selection, allocation, and rewarding of individual employees takes place on the basis of signals such as formal qualifications as well as on the basis of productivity. This is usually explained in terms of incomplete information and bounded rationality. The signals form a solution to this problem, as they are assumed to indicate the average productive capacities of the group to which they refer. The labor queue theory (Thurow, 1975) adds an interesting twist, pointing out that many relevant competencies are not even learned in education, but picked up through work experience on the job. According to this theory, education is an indicator of low training costs rather than high productivity. Finally, some scholars have questioned whether education has any effect at all on graduates' ability to perform, pointing out that this relationship is in fact weaker than that between education and reward (Bills, 2003). This has led credentialists such as Collins (1979) to claim that higher education does not lead to superior competencies but is used by "gatekeepers" to legitimize the rationing of access to high-status, highly paid jobs.

In reality, there is probably an element of truth in all these theories. The crucial point then comes down to specifying the contexts under which one or the other mechanism prevails. The mechanisms are likely to differ according to the kind of job or position, labor market segment (private/public, economic sector), and country. In a study like PIAAC, we might expect large differences between the countries in the extent to which skills affect labor market outcomes relative to the effects of educational credentials. There is strong evidence that in countries characterized by a high degree of selectivity, stratification, and standardization, employers are more likely to select and reward employees on the basis of formal educational qualifications than in countries where education is less regulated (Müller & Shavit, 1998).

Many of the control variables that are needed to get unbiased estimates of the effects of skills on economic and social outcomes are comparable to the ones discussed above on the effect of education and training on skills development, although education and training will now be treated as control variables instead of the predictor of interest. As indicated above, the highest level attained in formal education is one of the strongest predictors of skills. This is not only interesting in its own right, as a skill predictor or reporting category, but will likely be a confounding variable for many of the issues that policymakers are trying to understand in the context of PIAAC. Level of education is also a strong predictor of economic and social outcomes, and although this is often assumed to reflect differences in skill levels between levels of education, the precise causal mechanism is still somewhat controversial (Are the effects all directly attributable to human capital, or do theories of signaling and credentialism also tell part of the story?).

In this respect it is not only important to register highest formal level (which can be translated into number of years of formal schooling), but also the number of additional years of schooling that did not result in a diploma (which can be calculated as the difference between the year in which one last left education without completion (B_Q03c) and the year in which one last successfully completed formal education (B_Q01c). This schooling should lead to additional skills, and if the human capital theory is correct, to better outcomes.

In addition to level of education, labor market studies show large and robust differences in economic outcomes between fields of study in tertiary and secondary vocational education. Arts and humanities and social sciences often perform poorly, while business and engineering studies often do better than average. From a policy point of view, it is important to establish whether these differences are due to differences in the supply of and/or the demand for the skills of the graduates of these programs, to signaling or credentialism, to individual preferences, or to other factors.

The variables related to the number and intensity of received training is not only relevant in predicting skills, but also in predicting economic outcomes. As indicated above for education, the precise mechanism is not known and the estimates of the returns to training are biased by heterogeneous selection into training. For example, some people might get training because they are expected to be promoted instead of the other way around. We have included control variables like firm size (D_Q06a-b) to control for this unobserved heterogeneity. Most of these control variables are the same as the ones we discussed above. Additionally, when estimating effects of education and skills on outcomes, it is important to control for factors relating to household composition (J_Q01), family formation, as indicated by marital/cohabitant status (J_Q02a), and number and age range of children (J_Q03a-d), and job characteristics such as employee/self-employed status (D_Q04), supervisory status (D_Q08a-b) and job tenure (D_Q05a).

Skills and other outcomes

There is good empirical evidence that education not only affects labor market outcomes but is also a strong predictor of outcomes in other life domains. The BQ includes indicators of family formation (J_Q02a, J_Q03a-d), health (I_Q08), voluntary work (I_Q05f), political efficacy (I_Q06a) and social trust (IQ07a-b). Education not only affects the individual outcomes in these domains but also affects social returns as a result of spillover effects. This is one of the reasons why policymakers are so interested in understanding these broader effects of education, because the social returns in terms of decreased costs for health and crime may well overwhelm the individual economic returns. The OECD recently published a report on the social outcomes of learning (Schuller and Desjardin, 2007), underpinning this need for investment in education to increase health and civic and social engagement.

As with the effects of education on labor market outcomes, the effects of education on other outcomes are still not completely understood. Broadly, two mechanisms can be distinguished: an effect on skills and an effect on allocation. For the first effect we assume that education directly affects knowledge and skills that are relevant for healthy behavior, civic engagement, and so on. For instance, health programs may increase the knowledge of students in this area, leading to healthier behavior. The second mechanism refers to the role of education in allocating students to particular jobs or roles in society, for example, higher education increases the chance of ending up in healthier jobs or in social networks in which civic engagement is higher. In that case, the role of education is more indirect and it is not certain that investing in education will always have the anticipated effect. This is dependent on whether these outcomes are scarce resources or not. If people have to compete for scarce resources (as in the case of high-level jobs), investment in education changes the relative distribution but not the absolute.

From a policy point of view, it is therefore important to gain further insight into the underlying mechanisms. Moreover it is important to investigate to what extent low skills as a risk factor for social outcomes may be compensated for by other protective factors like job conditions, educational attainment, and so on.

3.3 The development of the JRA

In 2004 the OECD launched an initiative to develop a module in PIAAC on generic work skills requirements as a complement to the direct assessments. This was called the Job Requirements Approach (JRA). In the JRA, workers are asked to indicate the level of skills that is required in their current work in several skill domains. The basic idea of asking workers to report on skill requirements in their job is already older and has been successfully applied in different surveys, such as the British Skills Survey, similar surveys in Italy and Spain, the US O*NET survey, and several international graduate surveys (CHEERS and REFLEX).

The main arguments for developing a separate JRA module for PIAAC were the following:

- The direct assessments in PIAAC are limited to relatively few, albeit crucial, skill domains. Yet there was a widespread feeling, supported by some case studies, that other skills were becoming increasingly relevant in modern workplaces. Important examples were communication skills and the skills needed to work within teams, to work at multiple and flexible tasks, and to work more independently. There was also evidence that some of these skills, like computing skills, were being rewarded in the labor market over and above the returns to the education that people had received (Dickerson & Green, 2004). It was intended that the JRA module would provide a cost-effective way of assessing the relevance of these skills.

- Earlier skills surveys like IALS and ALL were mainly limited to the supply side of skills, that is, the stock of skills of the population. It was felt that some information on the demand side for skills was needed as well, that is, on the utilization of skills in the workplace. Sociological theory makes a distinction between “own skills” (the skills that individuals have) and "job skills" (the skills defined by jobs), and it was decided to measure some important job skills directly.

In the JRA module, respondents were asked questions about the skills that they use at work. First, the module generated many items describing the generic activities involved in doing the job. The choice of items was informed by theories of skill and the practices of commercial occupational psychology. To reduce the multiple items to a smaller and theoretically meaningful set of generic skills, statistical techniques were used to generate several generic skill indicators from the responses on these items.

In the course of development of the BQ, it became apparent that parts of the JRA module corresponded to a large degree to measures of skill use that are required for analyzing the results of the direct assessment. The subject matter expert groups (SMEGs) in the areas of literacy, numeracy and ICT developed scales that integrated the experiences from ALL with the newly developed insights from JRA. Scales were developed that measure the use of skills both at work and in everyday life (including study) in a similar way. These scales are broadly comparable to what has been measured in ALL, but the scales were adjusted to have better psychometric properties. Items are now included for the three central domains covered by the direct assessments literacy (reading: G_Q01a-h, H_Q01a-h; writing: G_Q02a-d, H_Q02a-d); numeracy (G_Q03a-h, H_Q03a-h); and ICT (G_Q04, G_Q05a-h, G_Q06-8, H_Q04a-b, H_Q05a-h).

In addition to these three central domains covered by the direct assessments, the JRA module contains items pertaining to problem solving (F_Q05a-b) as well as a range of interaction/social skills: cooperation (F_Q01b), influence (F_Q04a-b), managerial skills (F_Q03b), self-direction (F_Q03a, c), horizontal interaction (F_Q02a-c) and client interaction (F2d-e), and physical skills [stamina (F_Q06b) and manual skill (F_Q06c)].

Two assumptions underpin the use of the JRA. First, it is assumed that the individual is well-informed to report about the activities involved in the job he or she is doing. All jobs differ, even within quite narrowly categorized occupations, and one would normally expect the job-holder to know best. Nevertheless, this might not always be true, and where the job-holder has only been in a post for a short time, the assumption might be questioned. In the case of out-of-work respondents, the Field Test has assessed the reliability of respondents' ability to recall the activities of their most recent job in the previous 12 months. No indications were found that there was a serious recall bias. Second, it is assumed that the individual reports these activities in an unbiased way. This assumption might also be questioned: Individuals might talk up their jobs to boost their self-esteem. However, it is held that they are less likely to do so when reporting their activities than reporting how good they are in the performance of these activities. To minimize bias, the general principle was to ask respondents to report actual behavior, such as frequency of use and proportion of time spent on using different skills, rather than often-used alternatives such as the importance of these skills for the job.

The measures of "job skill" obtained through the JRA module are direct measures of the "own skill" held by respondents. Discrepancies between job holders' skills and job requirements are possible, however. Some individuals may have an excess supply of some skills and not be using them fully on the job; others may have insufficient skills for the job they are doing but may survive in the short run despite the consequent poor performance. These mismatches are dynamic: They can appear and disappear as both jobs and people change. In the domains that are also being directly tested, it will be possible to generate indicators of mismatch, where individuals have high levels of own skill and are in jobs where that same skill is used at a low level, or vice versa. There is also a general subjective question on self-perceived skill underutilization (F_Q07a). In several domains, however, there is no specific mismatch indicator available: The only indicator of skill in these domains will be the use of the skills in the job.

3.4 The development and validation of the BQ

3.4.1 The process of questionnaire development

Within the Consortium, the Research Centre for Education and the Labour Market was responsible for the development of the BQ. Advice on the BQ was given by the BQ expert group, consisting of the following members:

1. Prof. Ken Mayhew (chair), Pembroke College, Oxford and director of SKOPE, Research Centre on Skills, Knowledge and Organisational Performance.
2. Dr. Patrice deBroucker, Statistics Canada and member of OECD Network B.
3. Dr. Enrique Fernandez, European Foundation for the Improvement of Living and Working Conditions (Dublin, Ireland).
4. Prof. Francis Green, Professor of Labour Economics and Skills Development, Institute of Education, University of London
5. Prof. Masako Kurosawa, National Graduate Institute for Policy Studies, Japan.
6. Dr. Scott Murray, DataAngel Policy Research Incorporated.
7. Prof Jürgen Schupp, Honorary Professor for Sociology in the Faculty of Political and Social Sciences at the Free University and deputy director of the department Socioeconomic Panel Study at the German Institute for Economic Research DIW in Berlin.
8. Prof. Tom W. Smith, Director of the General Social Survey, National Opinion Research Center, University of Chicago.
9. Prof. Kea Tijdens, University of Amsterdam.
10. Prof. Robert Willis, Research Professor, Population Studies Center, University of Michigan.

Three meetings were held with the BQ expert group:

- 1-2 May 2008: Paris
- 23-24 June 2008: Offenbach (Frankfurt)
- 5-6 December 2010, Princeton, NJ, USA

Based on the discussions with the BQ expert group, several draft versions of the BQ were discussed with the NPMs, the BPC and the OECD.

For the inclusion of concepts and items in the BQ, we adopted the following list of criteria:

- The concepts must have a clearly established relation in the theoretical and empirical literature to skills and other relevant outcomes.
- Items must have good measurement properties in terms of reliability and validity and be able to maintain that over time.
- Items must be comparable across groups and across countries. This posed limits to items that may have been deemed vulnerable to cultural bias.
- Ex-ante harmonization was preferred over ex-post harmonization. National adaptations of questions (other than translation issues) were minimized and were only allowed in cases where it was functional (e.g., in asking about type of education, etc.).
- Wherever possible, items were preferred that were comparable with other international surveys. Most important was the comparability to IALS and ALL, but other international surveys such as the Labor Force Survey (LFS), World Value Survey (WVS) and the European Social Survey (ESS) constituted important markers as well.
- In general we recommended that most questions should be asked to everybody, or at least to a majority of the respondents. Developing items for small subgroups was minimized.

3.4.2 Two rounds of cognitive pre-tests

Rationale

Cognitive pretesting is an important tool for improving the quality and validity of questions (Willis, 2005; Beatty and Willis, 2007): They enable the identification of problems with the draft items, provide valuable insights into how the questions or specific terms are interpreted by respondents, how respondents use the given answer scales, how they recall (relevant) information, and how they make decisions and construct their responses. The results inform the evaluation and modification of survey questions.

As part of the overall validation strategy for the BQ, including the JRA, two subsequent rounds of cognitive pre-tests were therefore carried out with a selected subset of items.

The cognitive pre-tests were carried out in various countries and languages to forward PIAAC’s goal of achieving comparability of instrumentation and measures. Countries were chosen to represent a maximum bandwidth of cultural and language diversity. The first round of cognitive pre-tests was conducted from August to October 2008, the second round from October 2008 to January 2009. The pre-testing phase included item selection, translation and adaptation of these items, as well as the development and translation of an interview guide with general specifications for conducting the cognitive interviews including a scripted protocol. After the interviews were completed, both country-specific reports as well an overall report with combined findings and including recommendations were produced for both rounds.

The investigated questions were selected by an expert group identifying those items that (a) had not been tested and extensively used in previous studies and (b) appeared to be problematic in their formulations and/or response options. Given that the JRA items had already been validated in a separate pilot study, the two rounds of pre-tests focused on the feasibility of using the JRA for the recently unemployed.

Methodology

The following section shortly describes the item and country selection, the translation process, the specifications for the administration of the cognitive interviews, and the sample scheme.

Item selection

The item selection was based on version 3.1 of the BQ for the first round, and on version 3.5 of the BQ for the second round of cognitive pre-tests. Items were selected by staff from GESIS – Leibniz Institute for the Social Sciences, by experts in the field of cognitive pre-tests, and by item developers from the Research Centre for Education and the Labour Market (ROA). Items were selected according to criteria such as inclusion of crucial variables, inclusion of items with certain response categories and scales, or inclusion of items that had been identified as potentially problematic.

Due to the process of probing and follow-up probing, an item in the cognitive interviewing context requires much more time than in a standard interview. To reduce respondent burden, it was therefore necessary to limit the number of items and to restrict total interview duration, with 90 minutes recommended as an utmost maximum (Prüfer and Rexroth, 2005). Thus, for each of the two rounds of pre-tests, a total of 30 items were selected. Respondents answered different sets of questions depending on their education and employment status, with a maximum of 20 items per respondent.

Country selection

The first round of cognitive pre-tests was conducted in three PIAAC countries (United States, South Korea, and Germany), the second round in five PIAAC countries (United States, South Korea, Germany, Sweden and Spain). The countries were selected to cover three important linguistic areas and cultural regions: North America, Central Europe and Asia. Furthermore, this selection allowed the English source questionnaire to be pre-tested, thus ensuring that the potential problems identified in the cognitive pre-tests were not only due to translation, but rather to general design issues.

Item translation

Item translation was accomplished via double translation by two independent translators, followed by reconciliation. Problems and questions that arose during the translation process were communicated to the item developers and ambiguities were clarified.

Interviewer guide and techniques

An interviewer guide was developed by the cognitive pre-testing experts at GESIS – Leibniz Institute for the Social Sciences. The protocols in the interviewer guide integrated two techniques typically used in cognitive pre-tests: *Paraphrasing* and *probing*. The *paraphrasing* technique asks respondents to reformulate the question in their own words. This method provides information on how the item is understood by the respondents and whether this interpretation matches the question intent. The *probing* technique comes into play after the respondent has answered a survey question and focuses on specific issues (e.g., how the item is understood, potential ambiguities or reasons for choosing a specific answer category).

The interviewer guide specified item-by-item instructions on how to conduct the cognitive interview. It included information on probes and additional questions, as well as specifications for the data format. The interviewer guides were translated and used in each country to ensure that the same techniques and procedures were used for specific questions across all countries.

Administration of cognitive pre-tests

The cognitive interviews were carried out face to face and were audio-recorded. Prior to the cognitive pre-test, respondents were informed that the aim of the interview was to evaluate and improve questionnaire items, and not test the respondents. All institutes carrying out the cognitive pre-tests gave monetary incentives for participation.

Sample and quota scheme

The requested sample size was 25 respondents per country for each round of cognitive pre-tests, with a predefined quota scheme. This scheme called for respondents with specific combinations of education, and employment status, with a heterogeneous distribution of age and gender (cp. Table 3.1).

Table 3.1: Quota Scheme for Cognitive Pre-tests (Round 1 and Round 2) (N = 25)

	Lower educational level (ISCED < 3) N = 17		**Higher educational level (ISCED ≥ 3) N = 8**		
	Not in education	**Currently in education**	**Not in education**	**Currently in education**	**Total**
Job	6	---	3	---	**9**
Recent job	6	---	3	---	**9**
No recent job	3	2	1	1	**7**
Total	**15**	**2**	**7**	**1**	**25**

Results

All respondent-level information was carefully reviewed. This included evaluating the detailed protocol results, concrete responses to the items (e.g., which category on the answer scale), as well as spontaneous respondent reactions. Respondent results were supplemented by interviewer comments. As a result, item-specific recommendations were derived. Collating and merging results from interviewers and respondents from different cultural and linguistic backgrounds greatly enriched the pre-testing findings.

The results of the cognitive pre-tests Round 1 were presented at the NPM Meeting in Lisbon in October 2008 and a report was sent to participating countries, the item developers and the BQ expert group. The results of Round 2 were communicated through a written report in February 2009. The recommendations in the reports were considered for the further development of the BQ.

3.4.3 Analysis of Field Test data

The BQ for the PIAAC project was developed with a view towards supporting the three broad policy questions described above that are central to PIAAC as a whole. First of all, it was designed to provide a clear view of how skills are distributed in the adult population. The second broad policy question underpinning the PIAAC project was to establish why skills are important. The third was the need to determine what factors are related to skill acquisition and decline. It is these policy considerations that have shaped the selection of items for the BQ as used in the Field Test.

The analysis of the data from the Field Test was guided by a number of main goals. Regarding contents, it was primarily aimed at validating the BQ by examining its general feasibility, empirical item and scale properties, quality of the underlying concepts and its operationalization. Regarding length, it was first necessary to assess the average time needed to complete the questionnaire, or subsets of items, in order to estimate by how much the questionnaire needed to be reduced to achieve a practicable questionnaire length for the Main Study. The combination of these two analyses helped identify items that could potentially be removed while making sure that main reporting variables were retained and the BQ still addressed PIAAC's main policy goals.

Moreover, the analysis aimed at discovering irregularities in the country data sets that could reveal potential translation errors or technical problems during the BQ administration.

In order to achieve these goals, the following analyses were conducted:

- A timing analysis assessing the average duration of the administration of the BQ
- An item-based analysis focusing on item nonresponse, item response distribution and response duration
- A scale-based analysis assessing the reliability and functioning of the BQ's multi-item scales both within and across countries
- An analysis of the functioning of the items representing main concepts such as education and training, labor market and other outcomes, and noncognitive skills
- Routing checks of crucial filters and branching rules within the national BQs

All analyses were conducted at an overall (international) level. The timing analyses, the item-based analyses and the routing checks were also run at the country level. Countries included in the international item-based analysis and routing checks were Austria, Chile, Cyprus,[1] the Czech

[1] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Republic, Denmark, Estonia, Finland, Flanders (Belgium), England/Northern Ireland (UK), France, Germany, Ireland, Italy, Japan, Korea, the Netherlands, Norway, Poland, Portugal, Spain and Sweden.[2] At the time the analysis at the overall level was conducted (November 2010), the Field Test had not been completed in Canada, Slovakia, the Russian Federation[3] and the United States; they were thus not included in the overall analysis. In addition, national reports giving detailed information about item distributions, durations, routing and potential irregularities were provided to each country.[4]

The analyses of the multi-item scales and the functioning of the main concepts are based on PIAAC Field Test data of 18 countries – those included in the overall analysis, excluding England/Northern Ireland (UK), Flanders (Belgium) and Norway.

Completed interviews and partial completes, were taken into account in the analysis. Across countries, a total of N=81,597 interviews (completes and partial completes) were analyzed.

For all timing analyses, only completed cases were included. In order to eliminate outliers at the item level, the data were trimmed by replacing all item time values beyond +/- 4 times the median of each item per country with the value of +/- 4 times the median. All 20 countries mentioned above with the exception of Spain were included for the timing analysis. [5]

The assessment of the questionnaire length

With respect to the BQ, the main goal of the Field Test was to finalize the instrument to be used in the Main Study, which in practice primarily meant a significant reduction in length. The Field Test intentionally included more items than were to be implemented for the Main Study. This total Field Test questionnaire was estimated to take some 55 to 60 minutes on average. To make the Field Test as realistic as possible in terms of total time of the interview, it was decided to use a random module design. All respondents got a core questionnaire and one of four modules: one with questions on the use of nonliteracy skills at work (section F), one with questions on skill use in reading and writing (first parts of sections G and H), one with questions on skill use in numeracy and ICT (second part of sections G and H), and one module with questions on noncognitive skills and noneconomic outcomes (section I). This approach was thought to bring back the total interview time for the Field Test to some 40 minutes.

As the preparation for the Main Study needed to start quite soon after the data collection for the Field Test, the Consortium, countries and the OECD agreed early in 2010 on a two-phase process for revising and adapting materials for the Main Study BQ. Phase I took place prior to

---

[2] In **Estonia**, approximately 15% of the interviews had been administered in Russian language, and 85% in Estonian. In the analysis, interviews conducted in Russian were not taken into account. In **Portugal**, due to an error in the random assignment of the BQ modules, certain sections of the BQ were omitted from analysis. **Australia** did not share its dataset due to data confidentiality reasons.

[3] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[4] Including Canada, Slovakia, Russia and the United States but with the exception of Australia.

[5] There were doubts as to the representativeness in the sample taken from the population for the Field Test in Spain. This gave rise to a lower than average response duration in that country. For this reason, Spain was excluded from the database for the timing analysis.

the analysis of the Field Test data (between January and September 2010), while Phase II utilized results obtained in the analysis of the Field Test data (November and December 2010).

In the case of the BQ, Phase I of the revision process began with the creation of an interim BQ. This revised version of the Field Test BQ was based on recommendations provided by the OECD/BPC, which identified a prioritized list of questions to be deleted and areas of the questionnaire where further reduction in the number of items could be made if supported by results from the Field Test once that data was available. The interim BQ was finalized in June 2010.

Based on the Field Test data, this interim BQ was estimated to take some 45-50 minutes on average, ranging from under 40 minutes for inactive respondents to just over 50 minutes for employed respondents.[6]

The second step of the revision was data driven, which meant that it could only be implemented after the data collection for the Field Test had been conducted and an international data file prepared. The requirement was to reduce the BQ in its final form to a length of 40 minutes for the common core, with a maximum of an additional five minutes allowed for countries to add any national questions they considered necessary for their own purposes. This meant that the interim BQ needed to be cut back another 5-10 minutes. The rationale for dropping items in this interim BQ was based on a thorough analysis of the functioning of individual items as well of the concepts that were made up of individual items (e.g., scales). In the next paragraphs we report the main findings of the Field Test analysis on which we based the decision to further reduce and finalize the interim BQ.

Individual item functioning: item distribution and item nonresponse

Item nonresponse

Item nonresponse was assessed across countries with a focus on (a) questions on the individual's income, as these questions are known to have high nonresponse rates, and (b) questions asking about past behavior (retrospective questions), in order to explore whether the JRA questions could be administered to the currently unemployed estimating the requirements of their past job. Item nonresponse was also investigated at the level of single countries and language groups within countries, as this might indicate potential country-specific translation errors or technical problems during the BQ administration.

For most of the BQ items and across all countries, nonresponse was very low (1% or less per question). However, some items showed higher nonresponse rates, such as the open-ended income questions:[7] 9% for employees [6% refused (RF), 3% don't know (DK)] and 26% for self-employed (12% RF, 14 % DK). However, the follow-up questions asking those who did not

[6] The PIAAC BQ is a highly adaptive instrument with a large variety of routings depending on education, labor force status, and other variables. As we used a random module design, it was not possible to simply add up the time spent on the different items in the Field Test. Thus, different methods were used to arrive at a reliable time estimate and the time was calculated for different types of respondents: employed, unemployed and inactive with accompanying assumptions on the share of these people routed into different questions (e.g., the share receiving training).

[7] (D_Q16a - D_Q18c2).

respond to the open-ended questions to report their income in broad income categories proved to be effective. Indeed, among employees, the total item nonresponse decreased from 9% to 5% and for self-employed from 26% to 11% after having presented the follow-up questions.

In order to find out whether the JRA questions and other job-related questions were more difficult for those currently unemployed (but with work experience in the last year) than for those currently in paid work, the respective nonresponse rates between the two groups were compared. Results show that most job-related retrospective questions did not have increased "don't know" or refusal rates among the currently unemployed.

Item response distribution

Response distribution was examined in order to identify items or response categories (a) that could potentially be deleted or (b) that reveal irregularities indicating potential translation errors or data entry or coding issues. In addition to an initial visual inspection of the item response distribution and the number of respondents per item, statistical key figures of interval- and ratio-scaled items such as the mean, standard deviation, skewness and kurtosis were analyzed.

Overall, it can be stated that nearly all items behaved in the expected way and only minor issues were discovered.

Most *ordinal and nominal* level items were distributed as expected, that is, they showed sufficient coverage in terms of frequencies across all response categories. No severe floor or ceiling effects were detected for any of these items. For the majority of the *open-ended questions* with manually entered numeric data, improbable values were identified in the answers. However, the total number of respondents with improbable and/or impossible responses was very small. Moreover, some of these outliers seem to be simply caused by technical problems. For the Main Study this led to some revisions in the minimum and maximum values that could be assigned.

Two items showing slight problems in this regard were the intensity of formal qualification and the last learning activity. Results indicated that some respondents may have had difficulties in assessing the time spent on their formal education or learning activities, particularly when judging this time in terms of hours. For example, when asked for the hours spent on their formal education, 10% of respondents said they didn't know. Other respondents indicated they had spent 24 hours per day on learning activities (1%, n= 57).[8] Due to these problems as well as the limited predictive power of these items, the Consortium, in consultation with the BQ expert group, simplified this set of questions for the Main Study.

Potential translation/comprehension issues were identified with respect to the interviewer instructions for items regarding the number of learning activities during the past year:[9] Interviewers were advised to "count related learning activities held on different days as a single episode." In nearly half of the countries investigated, the maximum number of reported learning activities was 50. This might indicate that respondents and/or interviewers thought the number of *lessons* (in units) was meant, and not the number of *courses*. These observations led to some modifications in the interviewer instructions.

[8] Formal education: B_Q09a; time spent on learning activities: B_Q19b.
[9] Question B_Q12b-B_Q12h.

Some **other irregularities in data entry and coding or technical issues** were identified. For example, national adaptations were not coded back into the international core variables for all countries.

Routing

As the PIAAC BQ is a highly adaptive instrument, it contained several routings depending on education, labor force status and other variables, such as computer use.[10] In order to investigate the functionality of the national BQs, various routing checks were conducted in the Field Test analyses. In detail, it was tested whether respondents did indeed arrive at the questions they were intended to receive (and no other questions) as defined by the BQ design. These routing checks focused on the crucial filters and branching rules in the national BQs and were not exhaustive. For each country, two general types of checks were run: a) within section filters, to a large extent focusing on the operative functioning of the derived variables, and b) between section routings, also taking the BQ random modules into account.

The analysis of the routing showed that no systematic routing issues were observed and generally the flow of the BQ worked as intended in all countries. Overall, routing checks within and between sections yielded only a few minor issues and affected only individual cases. Within-routing checks showed, for example, that in 0.1% of the cases, respondents did not receive any computer use question even though at least one of these items was to be received by all respondents. Across-section routing checks revealed only very few incidents where problems with transitions from one into another section occurred. These problems were most likely due to technical issues.

Multi-item scale functioning in sections F, G and H

Four sections of the BQ contain collections of items that can be regarded as multiple indicators of the same construct. Section F contains a collection of items around different types of nonliteracy skills used at work, while sections G and H contain collections of items around literacy skills. Section I contains scales that largely address inter-individual differences in terms of perseverance, learning strategies, locus of control, and others. This section will be discussed in the next paragraph.

These four sections were also the ones that were subject to pseudo-random assignment of respondents (rotation), so that analytic strategies had to take into account that not all variables in these sections have been observed with all other variables. Also, the fact that four randomly assigned rotations were used limited the sample size to about one-fourth of the realized sample size in each country. Taking these limitations into account, the analyses conducted with the Field Test data from these sections took the following two routes:

Exploratory route: Section F contains up to three items per skill domain such as communication, planning, advising, and others. The interrelations between skill domains in section F, and how these skills are related to different occupations, were expected to be of interest for reporting the main test data. Exploratory analyses were carried out using factor analytic techniques

[10] Questions G_Q04 (“Do/Did you use a computer in your Job/Last job?”), H_Q04a (“Have you ever used a computer?”), and H_Q04b (“Do you use a computer in your everyday life now/outside work?”)

(summarized below), while other analytic techniques such as latent class analyses and hierarchical multidimensional item response modeling were also explored.

Confirmatory route: Sections G, H, and I contain well defined and larger collections of items around topics and can thus be regarded as psychological or behavioral scales. In order to evaluate the functioning of the scales in these sections, reliability analyses and scale refinement, as well as predictive analyses using a proxy of the respondent's test score (the so-called ETS zlogit score) were conducted.

Exploratory analyses of section F

For the exploratory analysis of section F, data were pooled across countries. Data from 21 countries as available on October 29, 2010 were used in the analysis.

The Consortium ran a factor analysis with subsequent Promax rotation. Given the results obtained from these analyses, it can be expected that profiles of skills based on the items in section F can be formed. The results of these analyses suggested factors that could be referred to as 1) cooperation, 2) advising, selling and negotiation, 3) teaching and presenting, 4) planning, and 5) physical work. It should be possible for these results and factors to be further refined in the future based on within-country analysis, because the Main Study provides sufficient sample size for such analyses.

However, two items did not perform as expected and were therefore recommended to be removed. Item F_Q01a appeared somewhat ambiguous, while item F_Q06a was redundant, covering much the same meaning as F_Q06b (correlation between the two was >0.7), which performed somewhat better in other respects.

Confirmatory analyses of sections G and H:

The items in sections G and H are to a large extent aimed at a parallel assessment of self-reported literacy skills around reading, writing, numeracy and ICT. Each item in these two sections belongs to exactly one of these four skill domains. Therefore, confirmatory 1-factor models (to check item coding is working as expected) and reliability analyses were conducted for four scales each in sections G and H.

The reliability analyses were conducted by country and then aggregated across countries yielding the following results:

- G_Q01 (Skill Use Work – Literacy – Reading).
    - The scale consisted of eight items: G_Q01a, G_Q01b, G_Q01c, G_Q01d, G_Q01e, G_Q01f, G_Q01g and G_Q01h. Across countries, the average Cronbach's Alpha was .82 (SD =0.04). The reliability did not increase when leaving any of the items out.
- G_Q02 (Skill Use Work – Literacy – Writing).
    - The scale consisted of four items: G_Q02a, G_Q02b, G_Q02c, and G_Q02d. Across countries, the average Cronbach's Alpha was .63 (SD =0.07). The reliability increased to a value of .66 when leaving item G_Q02b out (SD = 0.09). In addition,

the average item-total correlation of item D_Q02b was lower than .3 (M = .24, SD = 0.07).

- G_Q03 (Skill Use Work – Numeracy).
  - The scale consisted of eight items: G_Q03a, G_Q03b, G_Q03c, G_Q03d, G_Q03e, G_Q03f, G_Q03g and G_Q03h. Across countries, the average Cronbach's Alpha was .83 (SD =0.02). The reliability increased to a value of .85 when leaving out item G_Q03a (SD = 0.02). However, the average item-test correlation of item G_Q03a was close to .3 (M = 0.29, SD = 0.06).
- G_Q05 (Skill Use Work – ICT – Internet and Computer).
  - The scale consisted of eight items: G_Q05a, G_Q05b, G_Q05c, G_Q05d, G_Q05e, G_Q05f, G_Q05g and G_Q05h. Across countries, the average Cronbach's Alpha was .81 (SD =0.03). The reliability increased to a value of .82 when leaving out item G_Q03g (SD = 0.03). In addition, the average item-test correlation of item G_Q05g was smaller than .3 (M = .25, SD =0 .10).
- H_Q01 (Skill Use Everyday Life – Literacy – Reading).
  - The scale consisted of eight items: H_Q01a, H_Q01b, H_Q01c, H_Q01d, H_Q01e, H_Q01f, H_Q01g and H_Q01h. Across countries, the average Cronbach's Alpha was .72 (SD =0.04). The reliability increased slightly to a value of .72 when leaving out item H_Q01e (SD = 0.06). In addition, the average item-test correlation of item H_Q01e was smaller than .3 (M = .28, SD = 0.07).
- H_Q02 (Skill Use Everyday Life – Literacy – Writing).
  - The scale consisted of four items: H_Q02a, H_Q02b, H_Q02c and H_Q02d. Across countries, the average Cronbach's Alpha was .51 (SD =0.12). The reliability did not increase when leaving an item out.
- H_Q03 (Skill Use Everyday Life – Numeracy).
  - The scale consisted of eight items: H_Q03a, H_Q03b, H_Q03c, H_Q03d, H_Q03e, H_Q03f, H_Q03g and H_Q03h. Across countries, the average Cronbach's Alpha was .84 (SD =0.02). The reliability did not increase when leaving any of the items out.
- H_Q05 (Skill Use Everyday Life – ICT-Internet and Computer).
  - The scale consisted of eight items: H_Q05a, H_Q05b, H_Q05c, H_Q05d, H_Q05e, H_Q05f, H_Q05g and H_Q05h. Across countries, the average Cronbach's Alpha was .75 (SD =0.04). The reliability increased to a value of .76 when removing item H_Q05h (SD = .03). However, the average item-test correlation of item H_Q05h was larger than .3 (M = .32, SD = 0.09).

Except for the writing skill scales G_Q02 and H_Q02, the reliabilities of the scales in sections G and H were quite satisfactory. Note that the writing scales with four items each were also the shortest scales in the literacy skill-use sections.

Predictive analyses were also conducted on the scales in section G and H. Predictive analyses were conducted by country and then summarized across the 21 countries. The correlations of self reported skill-use scales with the zlogit score were at a moderate level and consistent across scales as well as countries. Compared to other measures such as the ones collected in section I, the skill use correlations with zlogit were higher. Note that even the least reliable (and shortest) writing skill use scale on average correlated with the zlogit 0.256 for skill use at home and 0.269 for skill use at work. The good consistency of skill use scales in terms of reliability and predictive validity led us to believe that these scales would be among the most valuable predictors of outcomes in modeling and reporting of the Main Study data.

Functioning of concepts

In this part of the analysis, we looked at the functioning of the key concepts in the BQ. We looked at items related to respondents' socioeconomic background, education and training, their labor market outcomes, some possibly relevant noncognitive skills, and some other outcome measures. We used a range of methods of analysis, including univariate (inspection of frequency distributions), bivariate (relation with other relevant indicators), scaling (mutual correlation of sets of items) and multivariate (relation with outcome measures, controlling for other characteristics) methods.

Background, education and training

Socioeconomic background (J_Q06b-e, J_Q07b-e)

The BQ contained five indicators of respondents' socioeconomic background, namely the highest level of education (in three broad categories) ever attained by both parents, the occupational code of both parents when the respondent was age 16, and the number of books in the household when the respondent was age 16 (as indicator of the level of cultural capital in the parental home). With the exception of some possible minor measurement issues in some countries, which were referred back to the countries involved for checking and where necessary correction prior to the Main Study, these variables all performed well in the analyses. They showed plausible frequency distributions and were related in the expected way to each other and to respondents' education, occupation, earnings and skills. This applied not just to bivariate relationships between the indicators of socioeconomic background and these characteristics of respondents, but continued to hold in multivariate analyses with controls for gender, age, field of study of highest completed education, employment status, immigrant status, cohabitation status, parenthood, country of residence and respondents own education and occupation (the latter with the exception of those analyses were these were the dependent variables).

However, the predictive power of parents' education was in almost all cases greater than that of parents' occupation, which added little additional explained variance once parents' education was included in the analyses. The only exception was when the respondent's own occupation was the dependent variable. Understandably, parents' occupation was in this case a better predictor than parents' education, but even here parent's education showed a significant effect.

Taking into account the length of time required for administering the questions on parents' occupation (around 1.5 minutes on average), it was decided that this was a strong candidate to be dropped from the BQ for the Main Study. The number of book in the parental household was a strong predictor of test score proxy and other relevant outcomes.

The recommendation was to retain items on parents' education and number of books in the household at age 16, but drop items on parents' occupation for the Main Study.

Level of education (B_Q01a, B_Q01a3, B_Q02b, B_Q02b3, B_Q03b, B_Q03b3, B_Q05a, B_Q05a3)

The component variables for this set of indicators were the highest completed level, the education level engaged in by those currently in education, and the highest level of education of programs that respondents may have started but failed to complete. All three indicators were asked separately for home country and foreign qualifications, so it was necessary to combine these into a single measure. All three indicators were initially composed of detailed ISCED codes distinguishing 13 levels as well as a category of "No formal qualification or below ISCED 1." For the purposes of most analyses this was recoded into three broad levels ("ISCED 1, 2 and 3C short," "ISCED 3C long, 3A-B and 4" and "ISCED 5 and 6)." Again with the exception of some minor country-specific issues, these variables performed well in the analyses. They were plausibly related to each other, as well as to respondents' occupation, earnings and skill level. Being currently engaged in education at a higher level than the highest completed level or having left education at a higher level without completion was associated with higher skill levels even after controlling for highest completed level of education.

Due to the extremely tight timeline available for revising the BQ for the Main Study, the separate items on level of foreign qualification for current, unfinished and recent education were already dropped prior to the analysis of the Field Test data. Because few respondents reported foreign qualifications, the data analysis provided no reason to reverse this decision.

Taking into account the fact that the separate items on level of foreign qualification for current, unfinished and recent education were already dropped prior to the data analysis, the recommendation was to retain the remaining set of items unchanged for the Main Study.

Field of study (B_Q01b, B_Q02c, B_Q05b)

For highest completed education, current education and other education followed in the last 12 months, respondents were asked to report their field of study (ISCED 97 broad fields of education and training, i.e., 1-digit codes). Apart from some minor country-specific issues, these variables all performed well in the analyses. They behaved well in terms of their frequency distributions, which were plausible and similar in all three cases, with a main exception that current and recent education tended less often to be general programs than highest completed education. This latter finding is consistent with the tendency for education to become progressively more specific as the educational career progresses.

In all fields of highest completed education, the most frequent choice of subsequent field of education was the same one. In addition to this relation with the field of study for current or recent education, the field of study of the highest completed education showed a clear and plausible relation with occupation, economic sector, gender, earnings and skill levels. This result held not only in bivariate analyses but in multivariate analyses that controlled for level of education as well as other relevant indicators such as gender, age, employment status, immigrant status, and country of residence. This confirmed that field of study is a relevant dimension in addition to the level of education.

The recommendation was to retain this full set of items unchanged for the Main Study.

Training participation and intensity (B_Q06-B_Q09, B_Q17-B_Q20, B_Q22-B_Q25)

Component variables for training participation and intensity were the number of training episodes in the last 12 months, hours of training of most recent episode, hours of training of second-most recent episode, proxy total time spent on training (a construct based on the former three variables), and the time spent in the last 12 months on education. These variables are inherently skewed: Most people follow little or no training, but a small number invest heavily in training. The skewedness is accentuated by some apparent measurement difficulties.

Several factors are likely to have contributed to these measurement difficulties. For the number of training episodes, it seemed likely that a small number of respondents reported repeated sessions of the same training episode (for example, a weekly language course) as separate episodes, which resulted in an implausibly large number of reported episodes for a small number of respondents. For hours of training in the two most recent episodes, there were also some implausibly high values, which seemed to be largely – although possibly not entirely – due to the fact that for those who opted to report training in weeks or days as opposed to hours, the final measure was based on answers to two separate questions that then needed to be multiplied with each other to produce the final measure. An error in either answer would therefore be multiplied and will thus result in an even larger error in the final indicator.

This problem was even greater for the proxy for total time spent on training, which was based on the number of training episodes and the time spent on the last two episodes. Because we lacked data on time spent on all but the last two episodes, this indicator was necessarily inaccurate at the individual level, and this problem was compounded by measurement error.

A more “holistic” method of measuring training duration was introduced for the Main Study, and which is believed to have reduced the measurement error, although it may not have removed it altogether because of the inherent difficulty of asking respondents to report the duration of all episodes combined. For the purposes of the analyses reported here, we assumed that high values on all the indicators were likely to be inaccurate, so we removed extreme values prior to analysis. After these adjustments, especially the indicator for training frequency was well behaved. The frequency distributions of all these variables appeared plausible. When related to other relevant indicators especially training frequency was well behaved, showing clear relations with level of education, occupation, earnings and skill level. These relations held up well in multivariate analyses after controlling for other relevant indicators, and training frequency and even training incidence (training yes/no) were also strong predictors of labor force status as well as noneconomic outcomes such as health, civic engagement and social trust. Training duration also showed some effects on other variables, but these effects were generally much weaker and less consistent.

Taking into account measurement issues, the limited predictive power as well as the length of time required for administering the questions on training duration (estimated at three minutes on average for the "holistic" measure of training duration proposed for the Main Study and an additional 1.5 minutes on average for the duration of participation in education in the last year), it was decided that these were strong candidates to be dropped from the BQ for the Main Study.

The recommendation was to retain questions on training frequency, but drop questions on training duration for the Main Study.

Labor market outcomes

Labor force status (C_Q01-C_Q05, C_Q07)

Formal labor force status, which differentiates the statuses "employed," "unemployed" and "not in the labor force," is constructed on the basis of answers to a series of questions on whether respondents are currently employed, available for work, waiting to start work, or have taken active steps to find work. There are two versions of this indicator, an Australian version automatically generated while the BQ is administered, and a European version. The difference between these two versions is both conceptually and empirically minor, with the sole difference being whether looking at job advertisements in the newspapers is regarded as an active step or not. There are only marginal differences in the frequency distribution in either case (several tenths of a percent shifting between "unemployed" and "not in the labor force"), and regardless of which version is used, these variables all performed well in the analyses. In both cases the frequency distribution was plausible. There was a clear relation between formal labor force status and subjective employment status (i.e., how respondents see themselves), but these were far from identical. However, the differences between subjective and objective status were plausible, with, for example, a considerable proportion of those who saw themselves as unemployed being formally out of the labor force. Labor force status was also related in a plausible fashion to education and skills. There was no real relation with parents' education, but this did not seem to indicate a problem with either indicator.

The recommendation was to retain this full set of items unchanged for the Main Study.

Earnings (D_Q16-D_Q18)

The gross earnings of respondents were measured by way of a separate set of questions asked to salaried and self-employed respondents. Respondents who were unable or reluctant to report precise earnings were given the opportunity to report earnings in broad categories. Salaried respondents were given the choice of reporting earnings per hour, day, week, two weeks or year, and were also asked to report any annual payments they received in addition to their regular pay package. Self-employed respondents who had conducted their own business for at least a year were asked to report their gross earnings from their business in the last year, and those who had conducted their business for less than a year were asked to report their earnings for the last month. Here as well, respondents who did not report precise earnings were given the opportunity to report in broad categories.

Based on assumptions on the earnings distribution and taking into account the basis on which salaried employees reported their earnings, the answers to all these questions were combined into an overall measure of hourly and monthly earnings, with a separate measure for salaried employees and self-employed as well as a combined measure for all respondents in paid employment. A thorough validation of earnings of the self-employed was not really feasible due to idiosyncrasies inherent in earnings from business (for example, many respondents reported zero earnings). The analyses presented here are based on earnings from salaried employment. The complex method of measuring earnings leads to some apparent measurement error for salaried workers. The causes of these problems are familiar from other research, with, for example, some respondents reporting hours worked in the last week rather than in a typical working week, but subsequently reporting typical earnings rather than the earnings corresponding to the reported hours. Because the final earnings indicator adjusts for hours worked, the resulting indicator will be flawed in cases when the reported hours deviate strongly from typical hours. For this reason we removed the top and bottom 2.5% of the distributions prior to the analyses presented here.

As anticipated, the use of broad categories as alternative to precise earnings was the exception rather than rule, but the inclusion of this option significantly reduced item nonresponse on earnings variables. After removal of extreme values, these variables all performed well in the analyses. The earnings distributions were still slightly skewed, but plausible. There were similar distributions in each country, with some variation in kurtosis and skewness. The broad categories worked very well, showing a highly similar distribution to directly reported earnings. Earnings were plausibly related to skills and to investments in training, as well as to respondents own education, and to parents' education and occupation.

The recommendation was to retain this full set of items unchanged for the Main Study.

Noncognitive skills

GRIT and locus of control (I_Q01-I_Q02)

The items under consideration here are related to three broad concepts: GRIT, self-discipline and locus of control. GRIT can be further subdivided into perseverance and consistency of effort, and locus of control into internal and external. None of these sets of variables formed a good scale, but internal locus of control achieved a Cronbach's alpha of 0.66, which is satisfactory for a scale consisting of only three items (the other alphas were: perseverance of effort, 0.53; consistency of effort, 0.54; GRIT combined scale, 0.59; self-discipline, 0.47; and external locus of control, 0.41). Although GRIT showed some relation to level of education and labor market outcomes, neither GRIT nor its subscales were convincingly related to test scores, and the bivariate relation with earnings disappeared in the multivariate analyses. For this reason, the Consortium recommended dropping all these items. Much the same applies to self-discipline, which in multivariate analyses was not related to test scores or economic outcomes. Both internal and external locus of control showed a clear bivariate relation with test scores, although only the effect of external locus of control held up in multivariate analyses. By contrast, internal locus of control showed clear effects in multivariate analyses of labor market outcomes. Closer inspection of the data revealed it was possible to develop a combined measure comprising two internal locus-of-control items and one (reversed) external locus-of-control item (representing roughly the concept of decisiveness) which performed well in multivariate analyses both on outcomes and test scores. However, the Consortium did not think this warranted keeping these items for the Main Study.

The recommendation was to drop all items related to GRIT, self-discipline and internal locus of control.

Time preference (I_Q03a-d)

This set of four items was dropped on the basis of the list of priorities provided by the OECD. The goal of analyzing this set of items was to establish whether that decision was justified, or whether strong reasons existed to reverse that decision. The analyses showed that, although this set of items formed an unreliable scale (Cronbach's alpha = 0.44), this scale performed surprisingly well in the multivariate analyses, showing among other things a strong positive relation to test scores, and for males also a clear relation with employment status and earnings. However, in the view of the Consortium, these results were not sufficient to warrant overturning the earlier decision.

The recommendation was to stand by the original decision to drop these items.

Learning strategy (I_Q04a-m)

This long set of items was intended to represent two related concepts: deep or elaborate learning and surface-rational learning. The results show we could form a reliable scale for deep or elaborate learning consisting of the following items: I_Q04b, I_Q04d, I_Q04h, I_Q04j, I_Q04l, I_Q04m (Cronbach's alpha = 0.78). We could not form a reliable scale for surface-rational learning. As the intention was to at least significantly reduce the number of these items retained for the Main Study, we therefore proposed dropping the remaining items, and evaluating the performance of the deep learning scale in multivariate analyses. These analyses showed mixed results. Importantly, however, it showed a strong positive relation with test scores, and inclusion of this indicator as a control variable resulted in significant changes in the estimated effects of education and training variables on skills. Although there was no really robust relation with other outcomes, on balance the strong relation with test scores and its impact as control variable made this, in our view, a strong candidate to be retained for the Main Study, together with internal locus of control. The recommendation was to retain the reduced set of six items for the Main Study.

Table 3.2 shows the average correlation of the zlogit proxy with scales in section I.

Table 3.2: Average correlation of section I scales with zlogit (proxy of skills)

Item	Average	SD
I_Q01_mean About Yourself - Grit and Self-Discipline	0.015	0.077
I_Q02_mean About Yourself - Locus of Control	0.148	0.087
I_Q03_mean About Yourself - Time Preference	0.219	0.070
I_Q04_mean About Yourself - Learning Strategies	0.145	0.088
I_Q06_mean About Yourself - Political Efficacy	0.211	0.059
I_Q07_mean About Yourself - Social Trust	0.093	0.080
I_Q03_mean About Yourself - Time Preference	0.219	0.070

Table 3.3 summarizes the main results of the effects of noncognitive skill scales in the multivariate analyses:

Table 3.3: Significant effects of section I scales in multivariate analyses.

	males' labor force status		females' labor force status		males	females	
Scale:	**unem-ployed**	**non-active**	**unem-ployed**	**non-active**	**hourly wage**	**hourly wage**	**test scores**
I_Q01 Grit, subscale perseverance of effort		nnn		nnn			
I_Q01 Grit, subscale consistency of effort							
I_Q01 Grit, combined scale		nnn		nnn			nnn
I_Q01 Self-discipline							
I_Q02 Internal Locus of Control	n	nnn		n	p	p	
I_Q02 External Locus of Control							nn
I_Q03 Time Preference	nnn						ppp
I_Q04 Learning Strategies: deep learning	ppp	p	pp				ppp

ppp/nnn: positive/negative effect significantly different from 0.0 at 1% level

pp/nn: positive/negative effect significantly different from 0.0 at 5% level

p/n: positive/negative effect significantly different from 0.0 at 10% level

Other outcomes

Civic engagement (I_Q05a-h)

In the June revision of the BQ, this set of items was replaced by a single item on voluntary work. This decision was vindicated by an initial inspection of the data, which shows a strong correlation between the two separate items in this topic in the Field Test version of the BQ. Among other things, the data showed that civic engagement is positively related to test scores and that this relationship entirely accounted for the bivariate relation between civic engagement and level of education. There was also a significant relation between civic engagement and immigrant status, labor force status and health status. We therefore believed this was a useful outcome variable that should be retained for the Main Study.

The recommendation was to retain single item on voluntary work for the Main Study.

Political efficacy (I_Q06a-d)

No reliable scale could be formed for this set of items (Cronbach's alpha = 0.47). However, in consultation with representatives of the OECD, it was decided it would be valuable to retain a single item indicator in order to maintain a diversity of noneconomic outcomes. After consulting an expert on this topic, the Consortium recommended keeping the first item (I_Q06a), which was felt to best reflect the meaning of individual political efficacy.

The recommendation was to retain single item on individual political efficacy for the Main Study.

Social trust (I_Q07a-d)

Although no strong scale could be formed, the first two items (I_Q07a and I_Q07b) have worked very well in the past in other surveys and together achieve a Cronbach's alpha of 0.64. This reduced scale was positively related to test scores, as well as to employment status, training participation, and current participation in education. Unexpectedly, it was also negatively related to deep learning.

The recommendation was to retain reduced scale of two items on social trust for the Main Study.

Health (I_Q08, I_Q09 and I_Q10)

It was decided to drop I_Q09 from the BQ. The remaining items seemed to perform well. However, both the bivariate and multivariate analyses showed that the subjective health question (I_Q08) worked a little better than the objective question (I_Q10). Most importantly, there was a clear relation of subjective health with test scores, but no such relation with objective health. Both health indicators were related to level of education, labor force status, and training participation. On balance, taking into account the clear relation with test scores and also that it has been well validated in earlier research, we felt that the subjective health indicator was preferable to the objective indicator.

The recommendation was to retain the single item on subjective health status for the Main Study.

Summary and conclusions

The most important result that can be reported on the basis of the analyses of the Field Test data is that the Field Test BQ to a very large extent succeeded in collecting the necessary information on respondents across countries. In addition, some decisions to delete variables made during the first phase of the revision process were supported by the Field Test data. For the most part, the items that were deleted in Phase I did not perform as well in certain respects as items that were retained, for example, in terms of high item nonresponse, proportion of the population covered. or performance in data analyses.

Applying the criteria noted in the introduction to the results of the Field Test analyses, the Consortium recommended removing the following list of items for the Main Study version of the BQ:

JRA items

In line with the JRA pilot analyses, most of the items in this section performed well. However, two items did not perform as expected and were therefore recommended to be removed. Item F_Q01a appeared somewhat ambiguous, while item F_Q06a was redundant, covering much the same meaning as F_Q06b (correlation between the two was >0.7), which performed somewhat better in other respects.

Skill use at work and in everyday life

In general the skill use items performed very well. We recommended that most be retained for the Main Study. However, two numeracy items and one ICT item did not perform as expected for both work and everyday life contexts and were recommended to be removed. G_Q03a/HQ03a did not show any consistent relation to skills and did not scale well, especially in the work context (lowest item total correlation among the group of items scaled together). G_Q03e/H_Q03e was part of a redundant item pair together with G_Q03f/H_Q03f (correlation was about 0.7 in both cases). It was decided to retain just one of these two items and drop G_Q03e/H_Q03e. The same was true for G_Q05b/H_Q05b, which covered much the same meaning as item G_Q05c/HQ5c (correlation was above 0.6 in both cases).

Section I

Perseverance and self-discipline (I_Q01a-I_Q01i)

This set of items did not perform well in terms of predictive power (average correlation with "test scores" was 0.015 across countries) and was at least partially redundant with respect to the concept of deep learning strategies.

Surface learning (I_Q04a, I_Q04c, I_Q04e, I_Q04f, I_Q04g, I_Q04i, I_Q04k)

This set of items showed poor scaling properties. However, deep learning formed a good scale and performed better in multivariate analyses.

Political efficacy (I_Q06b-I_Q06d)

This set of items showed poor scaling properties. However, in the interest of retaining a selection of noneconomic outcomes, it was recommended that the first of these items (I_Q06a) be retained because it was considered to be the most appropriate indicator of this variable among the four (highest average item total 0.307 among the four items, and explained >5% of variance of the "test score" proxy).

Social trust (I_Q07c, I_Q07d)

The first two items of the intended four-item scale performed reasonably well in terms of scaling properties and was recommended for retention. These two items are also the well-researched/established way of measuring social trust. The second two items did not perform as well and were recommended to be dropped from the Main Study instrument.

Disability (I_Q10a-I_Q10b)

We have a subjective overall health measure that performs well, showing a strong relation with "test scores," among other variables. The specific disability-related measures did not perform as well in comparison (no significant relation with "test scores") and were recommended to be dropped. In addition, the distribution of responses differed substantially across countries for these variables. These two items also were more time consuming than expected.

Intensity of formal education (items B_Q06-B_Q09a,b)

Intensity of last activity (B_Q17-B_Q20a,b)

There were some measurement problems with these items, and, in particular, with the summary measure for total time spent on formal education as well as nonformal learning activities based on these items. As pointed out earlier, Field Test results indicated that some respondents may have had difficulties in judging the time spent on their formal education or learning activities. In the case of nonformal learning, the question that related to the number of activities a respondent engaged in performed substantially better in multivariate analyses (for example, intensity of training activities explains no additional variance in "test scores" after a simple dummy (training yes/no) has been included). In addition to the measurement considerations, the administration time for these items was excessively long (1.5 minutes in the case of formal education, and an estimated three minutes in the case of nonformal learning for those who take these questions).

Even though these variables have analytical importance, the Consortium proposed to drop them.

- Mother's or female guardian's occupation (J_Q06c-J_Q06e)
- Father's or male guardian's occupation (J_Q07c-J_Q07e)

There was evidence that the question on mother's and father's education performed better in terms of predictive power in multivariate analysis. The occupation variable did not provide substantial incremental predictive power compared to the education variable in these analyses. In addition, the items were quite time consuming (1.5 minutes) and require human coding of responses compared to the education variables. The Consortium therefore proposed to drop these items.

The recommended revisions led to the required reduction in time for the Main Study BQ of some 10 minutes compared to the interim BQ. The expected average interview time for the international core BQ was therefore under 40 minutes – not including any national extensions.

3.5 The content of the Main Study BQ

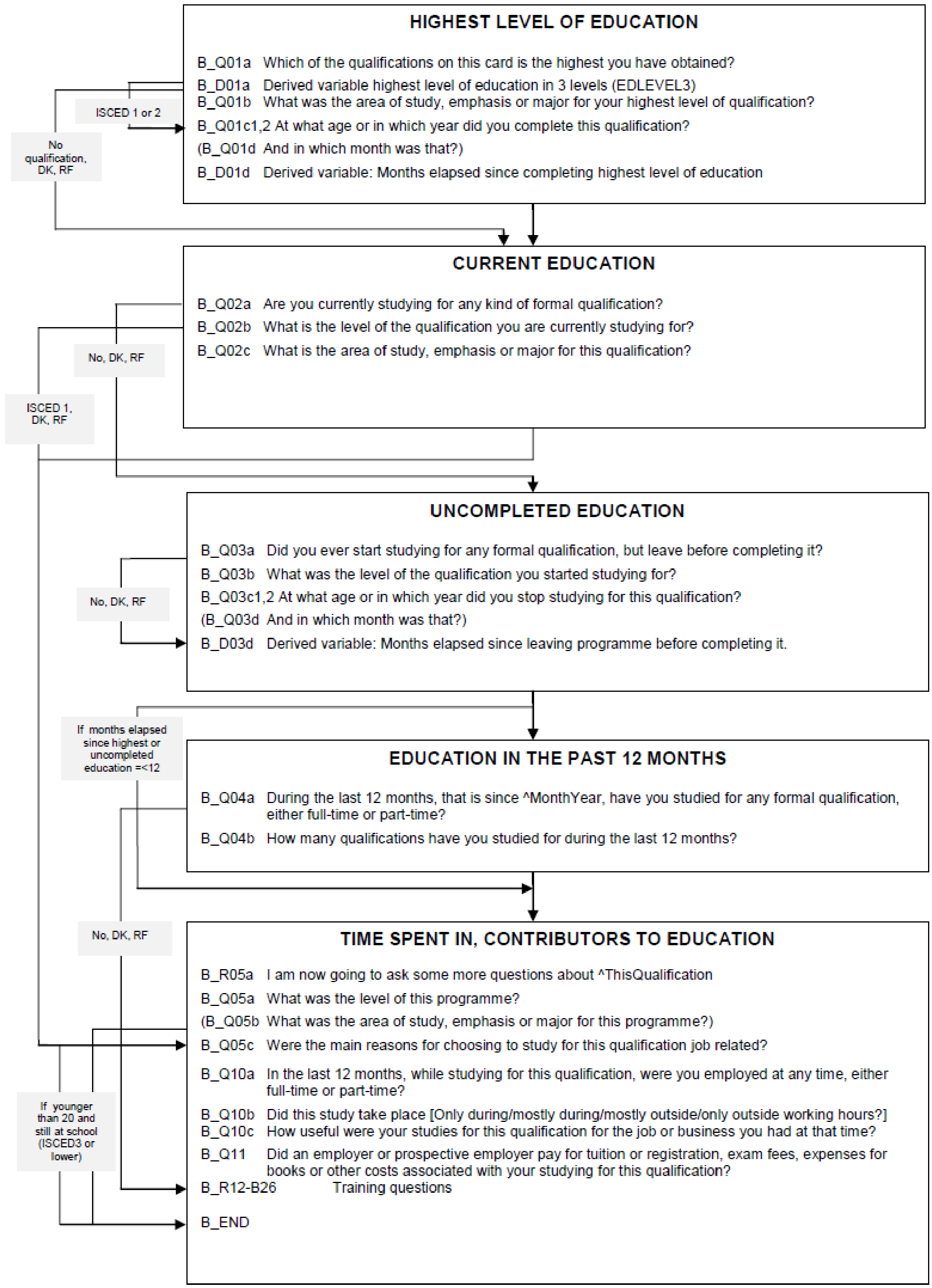
3.5.1 BQ Main Study

As indicated above, based on the analyses of the Field Test data, a final BQ for the Main Study was developed. The basic structure of the Main Study BQ is relatively straightforward, although some sections involve somewhat complex routing depending on, among other things, the educational and labor market status of the respondent. The BQ consists of a total of 10 sections:

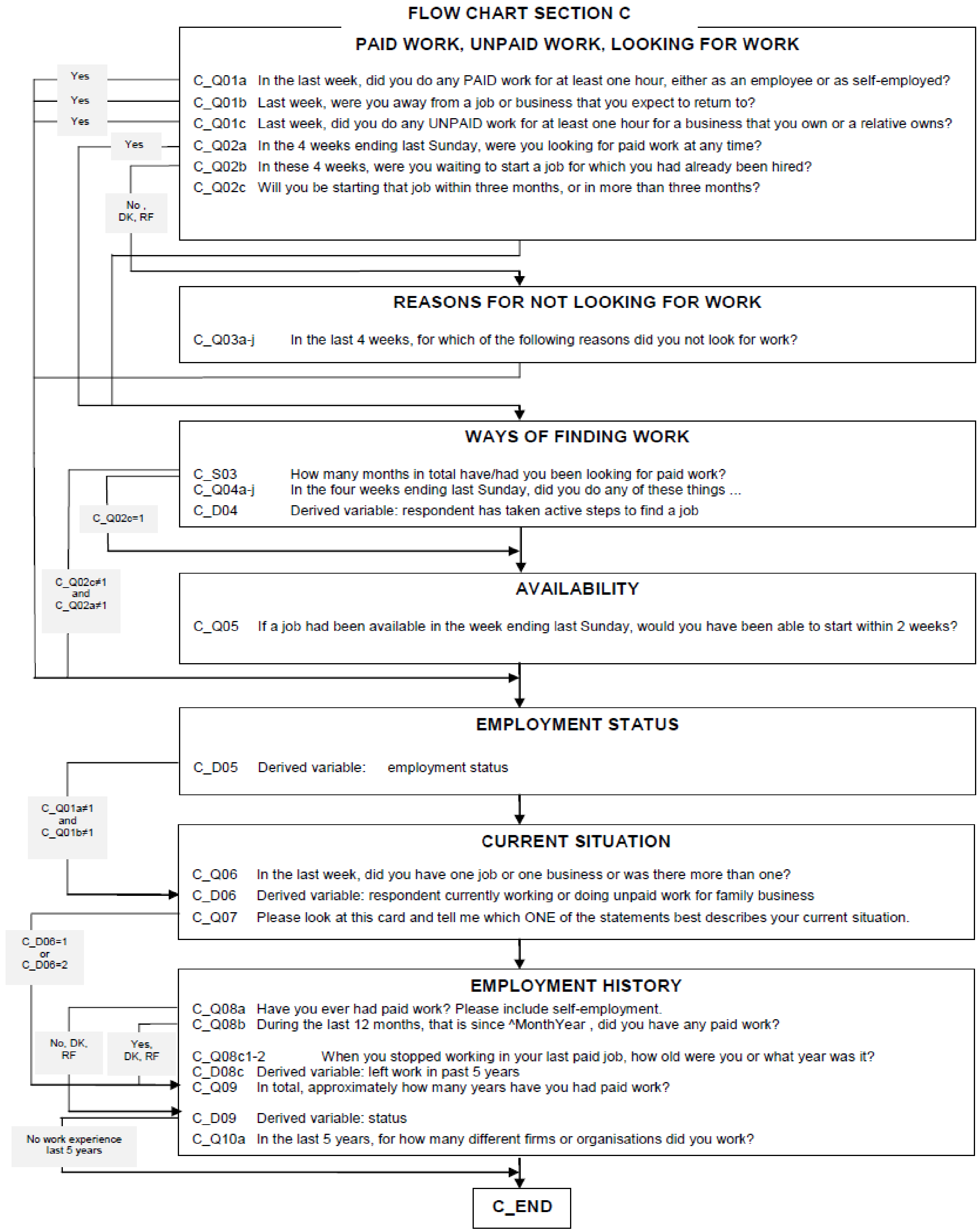
A. General information (date of birth, gender: all respondents)
B. Education and training (all respondents)
C. Current status and work history (all respondents)
D. Current work (for those currently employed or self-employed)
E. Last job (for those not currently employed or self-employed, who have worked in last five years)
F. Skills used at work (JRA Module; for those currently employed or employed in the last 12 months)
G. Skill use literacy, numeracy and ICT at work (for those currently employed or employed in the last 12 months)
H. Skill use literacy, numeracy and ICT in everyday life (all respondents)
I. About yourself (learning strategies, voluntary work, social trust, health: all respondents)
J. Background information (household composition, migration status, languages, parental education, cultural capital parental home: all respondents)

Sections B and C contain relatively complex routing. The following flow chart indicates the routing for section B.

FLOW CHART SECTION B



The following flow chart indicates the routing in section C.



3.5.2 National extensions

All countries were allowed limited scope to include national extensions they required for their own policy purposes. In order to avoid undue burden on respondents that could negatively affect the data quality, a strict rule was imposed that the total additional time added to the questionnaire in the form of such extensions was not allowed to exceed five minutes. The time estimates used to enforce this restriction took into account the number and type of proposed items to be added by a country.

3.5.3 National adaptations

The major adaptations countries were required to perform in the Main Study BQ were the following:

- Levels of education [highest, current, uncompleted and (other) recent education]: For obvious reasons, it was not feasible to use a standard international classification in order to ascertain the level of education a respondent is currently following or has followed in the past. In order to be comprehensible to respondents, all questions pertaining to level of education needed to be framed in terms of the qualifications currently or formerly available in the country concerned. Countries were required to develop an individual list of qualifications that could be directly matched to the standard list in the Master BQ to the extent that national equivalents for the levels described therein exist or have existed in the past. Countries were required to supply a full conversion scheme from their national levels to the international ISCED levels included in the Master BQ, including a specification of nominal years of schooling corresponding to each level and orientation and, where relevant, the vocational or academic nature of the program. A separate Excel sheet is provided with an overview of the national qualifications used in PIAAC with their conversion into ISCED level and orientation, nominal years of schooling, and vocational/academic.

- Country and language lists: Several questions in the BQ referred to countries or languages. These questions have a two-stage structure – first, a closed list comprising a limited number of countries/languages that are considered most relevant in the country concerned, and second, an open question for those respondents who wished to report a country/language not included in the standard list. Because the relevant countries and languages differ strongly from country to country, each country was required to adapt these items to national needs.

- In section C a block of questions was included that was designed to capture the respondent's job search behavior. Because search channels can differ subtly among countries, countries were asked to inspect the standard list of questions and, if necessary, adapt or add items to correspond to the national institutions and so on that may be involved.

- In sections D and E, several questions were used to ascertain the (last) occupation and (last) economic sector in which the respondent works or had worked in the past. Countries were required to check and, if necessary, adapt these items to the national setting.

- All income questions were asked in two forms: First, respondents were asked to report their income directly in the national currency. For respondents who were unwilling or unable to report directly their precise earnings, the option was made available to report in broad ranges. Countries were required to adapt the amounts and the currency used in these broad ranges based on explicit instructions how these should be derived from national statistics on recent population earnings distributions.

- For several questions throughout the BQ, countries were required to check, and if necessary adapt, the wording of questions to correctly reflect the national setting.

- Wherever adaptations involved some kind of structural change to the BQ – for example, splitting of a single item into multiple items, or the addition or deletion of one or more response categories – countries were required to make any necessary adaptations to routings, derived variables and so forth that make reference to the original items.

3.6 Quality check in the BQs

The BQ contained a number of features designed to assist the interviewers and ensure it was administered in a standardized way across all countries. These features were:

- Instructions given to interviewers. These instructions were designed to provide the interviewer with any relevant information that might be needed in order to pose the question in the correct manner, to indicate when to hand over and take back show cards, to provide support to respondents, and so on.

- Help buttons. In addition to these interviewer instructions, which were always visible to the interviewer but not read out to the respondent, the BQ contained a number of help buttons that the interviewer could consult if needed. These contained such things as additional information that could be provided to the respondent if needed, additional background information on the meaning or intent of questions, and so on.

- Consistency checks. For some items, there were consistency checks built in to the BQ that were triggered when a respondent gave a numeric answer to a question that might be considered to fall outside a plausible range of values. Examples include the age at which a respondent has reported a given event or status or the earnings reported by the respondent.

References

Adam, D., Bay, C., Bonsang, E., Germain, S., & Perelman S. (2006). *Occupational activities and cognitive reserve: A frontier approach applied to the survey on health, ageing, and retirement in Europe (SHARE)* (CREPP Report No. DP 2006-05). Retrieved from http://hdl.handle.net/2268/72515

Arrow, K .J. (1962). The economic implications of learning by doing. *Review of Economic Studies*, *24*, 155-173.

Arrow, K. J. (1973). Higher education as a filter. *Journal of political economy, 2(*33), 193-216.

Arthur, W. J., Bennett, W. J., Stanush, P. L., & McNelly T. L. (1998). Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance, 11(1)*, 57-101.

Bassanini, A., Booth, A. L. Booth, Brunello, G., De Paola, M., & Leuven, E. (2005). *Workplace Training in Europe*. (IZA Discussion Paper No. 1640). Retrieved from IZA website: http://ftp.iza.org/dp1640.pdf

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly, 71,* 287–311.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis with special reference to education*. New York: Columbia University Press.

Bills, D. (2003). Credentials, signals and screens: Explaining the relationship between schooling and job assignment. *Review of Educational Research, 73*, 441-70.

Binkley, M. R., Sternberg, R., Jones, S., Nohara, D., Murray, T. S. & Clermont Y. (2003). Moving towards measurement: The overarching conceptual framework for the ALL study. In T. S. Murray, Y. Clermont, & M. Binkley (Eds.), *Measuring adult literacy and life skills: New frameworks for assessment*. Ottawa: Statistics Canada.

Borghans, L. A., Duckworth, L., Heckman, J. J. & Ter Weel, B. (2007). *The economics and psychology of cognitive and non-cognitive traits* (working paper). Chicago: University of Chicago.

Borghans, L. A., Golsteyn, B., de Grip, A. (2006). *Meer werken is meer leren: Determinanten van kennisontwikkeling*. [English translation: Working more is learning more: determinants of knowledge development] 's-Hertogenbosch: CINOP.

Boudon, R. (1974). *Education, opportunity and social inequality.* New York: John Wiley & Sons.

Bourdieu, P. (1984). *Distinction: A social critique of the judgment of taste.* London: Routledge & Kegan Paul.

Brunello, G. (2004). On the complementarity between education and training in Europe. In D. Checchi and C. Lucifora (Eds.), *Education, training and labour market policies in Europe.* Retrieved from http://www.palgraveconnect.com/pc/doifinder/10.1057/9780230522657

Collins, R. (1979). *The credential society: An historical sociology of education and stratification.* New York: Academic Press.

De Corte, E. (1990). Towards powerful learning environments for the acquisition of problem-solving skills. *European Journal of Psychology of Education*, *5,* 5-19.

De Grip, A., Bosma, A. H., Willems, D., & van Boxtel, M. (in press). Job-worker mismatch and cognitive decline*.* *Oxford Economic Papers*.

De Grip, A., & Van Loo, J. (2002). The economics of skills obsolescence: A review. In A. de Grip, J. van Loo and K. Mayhew (Eds.), *The economics of skills obsolescence, research in labour economics, Vol. 21*, (pp. 1-26) Amsterdam/Boston: JAI Press.

Dickerson, A., & Green, F. (2004). The growth and valuation of computing and other generic skills. *Oxford Economic Papers-New Series, 56*(3), 371-406.

Glaser, R. (1991). The maturing of the relationship between science of learning and cognition and educational practice. *Learning and Instruction, 1,* 129-144.

Gruber, J., & Wise, D. (2004). *Social security programs and retirement around the world*. Chicago: University of Chicago Press.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labour market outcomes and social behavior. *Journal Of Labour Economics*, *24*(3), 411-482.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice Hall.

Müller, W., & Shavit, Y. (1998). The institutional embeddedness of the stratification process: A comparative study of qualifications and occupations in thirteen countries, In Y. Shavit & W. Müller, *From school to work. A comparative study of educational qualifications and occupational destinations*. Oxford: Clarendon Press.

Prüfer, P., & Rexroth, M. (2005). Kognitive Interviews. [Cognitive Interviewing]. *ZUMA How-To-Reihe,* 15. Mannheim: ZUMA.

Rychen, D. S., & Salganik, L. H. (Eds.) (2003). *Key competencies for a successful life and a well-functioning society*. Göttingen, Germany: Hogrefe & Huber.

Schooler, C., Mulatu, M. S., & Oates, G. (1999). The continuing effects of substantively complex work on the intellectual functioning of older workers. *Psychology and Aging, 14,* 483–506.

Schuller, T., & Desjardins, R. (2007). *Understanding the social outcomes of learning.* Paris: Organisation for Economic Co-operation and Development.

Schultz, T. (1963). *The economic value of education.* New York: Columbia University Press.

Spence, M. (1973). Job market signalling, *Quarterly Journal of Economics, 87*(1), 355-374.

Thurow, L. C. (1975). *Generating inequality*. New York: Basic Books.

Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganic (Eds.), *Defining and selecting key competencies* (pp. 45-65). Göttingen, Germany: Hogrefe & Huber.

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.

Chapter 4: Translation, Adaptation, and Verification of Test and Survey Materials

Andrea Ferrari and Laura Wayrynen, cApStAn; Dorothée Behr and Anouk Zabal, GESIS

4.1 Overview

The PIAAC assessment instruments (comprising cognitive instruments and the BQ) were originally developed in English, but administered to sampled adults in their own language. It follows that the successful *localization* of assessment instruments is an important component of the project. Some definitions, first, of terms which we shall use throughout this chapter:

- *Localization* can be defined, in general terms, as the process of adapting a product or service to a particular language and culture. A successfully localized product or service is one that appears to have been developed within the local culture. For international comparative assessment studies, like PIAAC, the challenge is to localize test and questionnaire items while maintaining the comparability of collected assessment results and contextual data across countries and languages.

- The *localization process* can be broken down into *translation/adaptation* and *validation*. The words *translation* and *adaptation* are used jointly because the term translation is deemed too restrictive to describe the process of culturally adjusting a test rather than literally translating it. An adaptation may consist in changing the picture of a stimulus, in changing the combination of July/summer to July/winter (or January/summer) for the Southern hemisphere, in changing a coeducational school context to a boys' or girls' school context for certain countries, etc. It may, e.g., involve a change of wording, register, context, currency, measurement unit, or form of address. *Validation* refers to quality control steps which will be defined later.

In PIAAC, as in many major international assessment studies, the localization process followed a mostly decentralized model:

- The participating countries (National Centers) were each responsible for localizing assessment materials for use in their respective countries.

- The PIAAC Consortium guided and assisted the countries throughout the process, in particular by developing and conducting linguistic quality assurance (LQA) and linguistic quality control (LQC) processes.

In PIAAC, the LQA processes implemented by cApStAn in cooperation with other Consortium players included:

- Early resolution of potential localization issues, via preliminary scrutiny of source assessment materials to anticipate adaptation issues, ambiguities, cultural issues, or item translatability problems.

- Definition of the localization design, based on the OECD PISA (Programme for International Student Assessment) design. The minimum standards to be followed by countries included a double translation and reconciliation design, making use of professional staff, and attending the training sessions organized by the Consortium. The key quality-control steps included in the design were the verification of National Centers' initial submissions by verifiers appointed, trained and monitored by cApStAn, a final check of instruments after post-verification revision by National Centers, and layout corrections by Consortium technical staff, and the documentation of all steps leading to the finalized localized instruments.

- Preparation of general translation and adaptation guidelines, separately for the BQ and the assessment materials. These key documents set out requirements and roles, translation traps, pointers on linguistic difficulty, psychometric traps, cultural adaptations, etc. They are further described in sections 4.3.1 and 4.3.2.

- Preparation of centralized tools for documenting and monitoring the successive translation, adaptation and verification activities: the VFFs (Verification Follow-up Forms) and BQAS (Background Questionnaire Adaptation Spreadsheets) used in the Field Test and later the MMFs (Main Study Translation-Adaptation-Verification Monitoring Forms). These tools included detailed item-specific translation and adaptation guidelines such as advice on adaptations that were mandatory, desirable or ruled out; advice on terminology problems and idiomatic expressions, literal or synonymous matches, that is, between stimuli and items to be echoed, patterns in response options to be echoed, formatting issues, and so on. Figure 4.1 shows an example of a VFF with item-specific guidelines

Figure 4.1: Example of a VFF

PIAAC
FIELD TRIAL 2009

VERIFICATION FOLLOW-UP FORM

COMPUTER-BASED

Country: PT
Target language: pt

UNIT: Election Results | PIAAC ID: C302BC02 | ALL ID: COREQ2S1

PLEASE INSERT NEW LINES, IF NEEDED, TO DOCUMENT ADDITIONAL ISSUES

LOCATION	ENGLISH SOURCE	PROPOSED TARGET VERSION	CONSORTIUM RECOMMENDATION	NPM COMMENT	VI INTE
stimulus	Nationwide Manufacturing Company Union Council ELECTION RESULTS		Note: 'Union' is to be understood as trade union, i.e. an organization representing workers		
stimulus	Posting Date: June 22~~, 2000~~		Eliminate ', 2000' versus ALL version		
stimulus	The election of a new member of the Union Council for election group 3, at the Carver plant took place on June 21~~, 2000~~.		The name 'Carver' may be changed. Note: 'plant' means here 'factory' Eliminate ', 2000' versus ALL version.		
stimulus	The results of the election were as follows:				
stimulus	Candidates Number of votes A. Greer 120 votes H.A. Holliday 80 votes G. F. Reynolds 29 votes		Names of people may be changed. Keep the three numbers aligned over each other.		
stimulus	Consequently Mr. A. Greer was formally elected as a member of the Union Council for Nationwide Manufacturing Company.		If name 'A. Greer' is changed, change it here too		
stimulus	In accordance with article 16, paragraph 1 of the Union Council bylaws, any interested party may lodge a complaint with the council within one week after publication of these results.				

- Provision of training sessions for countries' translation teams or their trainers of translations. A general session was provided at a meeting in Lisbon, Portugal, in October 2008, and modular workshops (for the various types of materials) were provided at a Barcelona, Spain, meeting in March 2009.

- Provision of a translation training kit so that further training sessions could be held in countries. The kit included a customizable PowerPoint presentation, materials for hands-on exercises, confidentiality forms, and so on.

- Continued assistance to National Centers throughout the localization process (help desk via ticketing system, see Chapter 6).

In PIAAC, the implemented LQC processes included:

- Verification by the Consortium of target versions submitted by National Centers against the source versions, with reporting of residual errors and undocumented deviations, and expert advice where corrective action was needed:

  - For Field Test instruments: full verification of all national materials

  - For Main Study instruments: "focused" verification of changes made by countries to their finalized Field Test national materials (whether to echo changes made to the source version or at the initiative of the National Centers), extra checks for risky cases as needed, and full verification of newly translated materials

- A final check procedure after National Centers carried out their post-verification revision of instruments and Consortium technical staff made layout corrections, again with reporting and follow-up of residual errors and/or unresolved issues.

- The scope of verification included all translated instruments viewed by respondents (computer-administered test units, help and orientations, BQ, paper test booklets) as well as language-dependent automated scoring rules (for the “highlight in stimulus” response mode and numeric entry response mode), paper scoring guides, and the “CAPI workflow” file used by interviewers to conduct the questionnaire and assessment sessions.

4.2 Participation in the development of the source version

Early resolution of potential localization issues via preliminary scrutiny of source assessment materials is an upstream LQA process which aims to reduce the difficulties and workload encountered later downstream. cApStAn reviewed the first drafts of new cognitive materials and of the BQ (as of version 3.4) with an eye to anticipating adaptation issues, ambiguities, cultural issues, or item translatability problems, with suggestions for either rewording or adding item-specific translation/adaptation guidelines.

cApStAn also provided English translations of item submissions from participating countries in Japanese, Italian, German and French; some of these were selected to be part of the PIAAC item pool.

Throughout the localization process, cApStAn took care of an errata management process, whereby errors in the source identified by National Centers or verifiers were tracked and, depending on the nature of the error and the time of discovery, listed for correction in source and/or national versions either at Field Test or Main Study phase.

4.3 Testing languages and translation/adaptation procedures, including double translation design

4.3.1 Testing languages and translation/adaptation procedures for the BQs

The major bulk of translation occurred in preparation for the Field Test. Therefore, the focus in the following will be on translation activities prior to the Field Test rather than in preparation of the Main Study.

The BQ was translated/adapted from (international) English into 39 national versions comprising 26 languages including English. Table 4.1 displays the languages of the BQ for each country.

Table 4.1: Languages of BQ for each country

Country	Languages
Australia	English
Austria	German, Turkish, Serbo-Croatian
Canada	English, French
Chile[1]	Spanish
Cyprus[2]	Greek
Czech Republic	Czech
Denmark	Danish
England/N. Ireland (UK)	English
Estonia	Estonian, Russian
Finland	Finnish, Swedish
Flanders (Belgium)	Dutch
France	French
Germany	German
Hungary[3]	Hungarian
Ireland	English
Italy	Italian
Japan	Japanese
Korea, Rep. of	Korean
Netherlands	Dutch
Norway	Norwegian (BM), English
Poland	Polish
Portugal[4]	Portuguese
Russian Fed.[5]	Russian
Slovak Rep.	Slovak, Hungarian
Spain	Spanish, Catalan, Galician, Valencian, Basque
Sweden	Swedish
United States	English, Spanish

[1] Chile later dropped out of this cycle of PIAAC and joined PIAAC Round 2.
[2] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[3] Hungary later dropped out of PIAAC.
[4] Portugal later dropped out of PIAAC.
[5] Please refer to the note regarding the Russian Federation in the Note to Readers section of this report.

Austria, Norway and the United States translated the BQ into more languages than they did for the assessment instruments. This was to accommodate important non-English speaking populations.

Prior to the BQ translation, each country, in cooperation with the Consortium, adapted the international BQ version to its local context. Adaptations at this stage mainly pertained to questions that, although measuring the same underlying concept, were in themselves substantively different from country to country (e.g., education, occupation or income items). Furthermore, countries were offered the opportunity to add items of country-specific interest not yet included in the international BQ. All such adaptations and national extensions were subject to approval by the Consortium. Chapter 3 describes the process of adaptation and extension in detail.

Once the adaptations and national extensions received signoff from the Consortium, a country-specific BQ version was built for each country, consisting of the (adapted) common set of international BQ items and the country-specific items. This version served as the basis for translation. The translation environments for the BQ translation were the Item Management Portal as well as specific translation software. Chapter 6 describes the technical tools.

Because comparability of survey materials is essential to any meaningful use of cross-national survey data, countries received a general guideline document laying down a quality framework for translation. The guideline document for the BQ translation focused on the one hand on the general translation process and on the other hand on issues to consider in the actual translation.

The guidelines on the general translation process included the recommended translation approach of double translation by two independent translators, followed by reconciliation. Double translation allows the spotting of misinterpretations or ambiguities, idiosyncratic wording or simply translator oversights; moreover, it offers stylistic variants among which to choose in light of a fluent translation. It has established itself as a state-of-the art approach in questionnaire translation. For reconciliation, team reconciliation was proposed to countries as a very efficient reconciliation method. Team reconciliation brings together at one table a unique mix of competencies: translators and linguistic experts, experts in the various domains of the questionnaire (education, work, etc.) as well as experts in questionnaire design and survey methodology. This broad range of expertise (translation, domain, design) is regarded as essential for producing high-quality questionnaire translations (Harkness, 2003; Harkness, Villar, & Edwards 2010). Alternatively, as a minimum, a single reconciler was required, ideally with input from a panel of experts in survey methodology and the various domains covered by the BQ. Translators were to be skilled practitioners, translating into their mother tongue and experienced or trained in questionnaire translation. Reconcilers were to have strong language skills in both source and target languages and be knowledgeable about questionnaire translation, questionnaire design, and the content domains covered by the BQ.

The general BQ guidelines also specified an overall framework for the BQ translation. The fact that a number of items in the BQ had been taken (changed or unchanged) from other surveys was acknowledged. Countries were given freedom to consult already existing translations from these surveys. However, it was stressed that in the end, the adherence and comparability to the PIAAC BQ was the crucial factor and would be the basis for verification.

The guidelines on issues to consider during translation specified that countries were to produce a questionnaire translation that maintains the measurement properties and the meaning of the source questionnaire, while at the same time being as fluent and understandable as possible. The overall task was to strike the right balance between faithfulness and fluency. The general message to countries was to produce the best possible *translation*. Any adaptations – beyond those already been agreed on – that countries deemed necessary had to be documented by countries and submitted to the Consortium for approval. Adaptations in this case were understood as intended deviations from the source version going beyond the changes that typically occur through translation. While the adaptations occurring *prior* to the translation phase applied to all countries in the same manner (e.g., all countries had to implement their own education measures), adaptations *during* the translation phase, if occurring at all, affected individual countries only. Countries were provided with an Excel tool in which to document adaptation needs: They were asked to provide an explanatory back translation of their chosen translation into English and to justify their decision. Back-translation in PIAAC was thus seen as a tool enabling communication with the Consortium and allowing for a commonly understood documentation. It did not serve as an assessment tool in itself.

Furthermore, countries were given item-specific translation guidelines. These provided further clarifications (e.g., on the meaning of terms or phrases or on characteristics of response categories) for a certain number of questionnaire items. The need for these clarifications had been identified by expert reviews focusing on potential translation problems and results of the cognitive pre-test. Furthermore, a so-called advance translation had been conducted on a pre-final version of the BQ (cf. Dorer, 2012). The goal of this translation was to identify problems in the source questionnaire while it was still under development and to take appropriate action (e.g., adding item-specific guidelines, changing wording).

During a one-day workshop at the NPM meeting in Barcelona in March 2009, NPMs, national staff responsible for the translation process, or translators themselves were introduced to the specificities of the translation of the BQ. The workshop covered the technical environment of the questionnaire translation (Item Management Portal, translation software), the different types of BQ translation guidelines, as well as good translation practice and discussion. The national teams were encouraged to replicate (parts of) the workshop in their countries with their chosen personnel.

During the translation and reconciliation process itself, countries were given the opportunity to ask queries about any problems they encountered (regarding meaning, technical issues, etc.). These queries were submitted by countries within the Open Ticket Request System (OTRS); GESIS monitored and answered the BQ questions and liaised with other Consortium partners as needed – in particular ROA as item developer of the BQ.

After reconciliation, the BQ translations were submitted by the countries to the Consortium along with any documentation on special translation decisions and desired adaptations. The BQ translations (and adaptations) underwent the same verification procedures as the assessment materials. Subsequent parts of this chapter present the verification process.

After the Field Test, countries had the opportunity to correct translation errors that had come to their attention in the course of fieldwork or their own analyses. Furthermore, the Consortium provided each country with a PIAAC Field Test Report which included recommendations to check certain specific items (where applicable). However, modifications to the questionnaire were required to be restricted to those that were absolutely necessary, i.e. to correct *errors*, but not make any changes such as stylistic improvements which could otherwise affect item functioning for items which had proved to work well in the Field Test.

4.3.2 Testing languages and translation/adaptation procedures for the cognitive instruments

The cognitive instruments were translated/adapted from the international English source version into 35 national versions comprising 24 languages, as shown in Table below, which includes information on participation in the two core components of literacy and numeracy as well as the two optional components of problem solving in technology-rich environments (PSTRE) and reading components. (Note that some countries translated the BQ into more languages than they did for the assessment instruments; this information is given in the previous section.)

Table 4.2: Translation by country for cognitive instruments

Country	Languages	Literacy/Numeracy	PSTRE	Reading Components
Australia	English	Yes	Yes	Yes
Austria	German	Yes	Yes	Yes
Canada	English, French	Yes	Yes	Yes
Chile[6]	Spanish	Yes	Yes	Yes
Cyprus[7]	Greek	Yes	NA	Yes
Czech Republic	Czech	Yes	Yes	Yes
Denmark	Danish	Yes	Yes	Yes
England/N. Ireland (UK)	English	Yes	Yes	Yes
Estonia	Estonian, Russian	Yes	Yes	Yes
Finland	Finnish, Swedish	Yes	Yes	NA
Flanders (Belgium)	Dutch	Yes	Yes	Yes
France	French	Yes	NA	NA
Germany	German	Yes	Yes	Yes
Hungary[8]	Hungarian	Yes	Yes	Yes
Ireland[9]	English	Yes	Yes	Yes
Italy	Italian	Yes	NA	Yes
Japan	Japanese	Yes	Yes	NA
Korea, Rep. of	Korean	Yes	Yes	Yes

[6] Chile later dropped out of this cycle of PIAAC and joined PIAAC Round 2.
[7] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[8] Hungary later dropped out of PIAAC.
[9] Ireland joined late but was able to borrow and adapt the UK English version

Country	Languages	Literacy/Numeracy	PSTRE	Reading Components
Netherlands	Dutch	Yes	Yes	Yes
Norway	Norwegian (BM)	Yes	Yes	Yes
Poland	Polish	Yes	Yes	Yes
Portugal[10]	Portuguese	Yes	Yes	Yes
Russian Fed.[11]	Russian	Yes	Yes	NA
Slovak Rep.	Slovak, Hungarian	Yes	Yes	Yes
Spain	Spanish, Catalan, Galician, Valencian, Basque	Yes	NA	Yes
Sweden	Swedish	Yes	Yes	Yes
United States	English	Yes	Yes	Yes

The translation environment for the cognitive instruments was the same as for the BQ translation: the OLT (Open Language Tool) translation software used for XLIFF files exchanged via the PIAAC Item Management Portal, described in detail in Chapter 6. XLIFF is the abbreviation of XML Localization Interchange File Format – a standard file format which permits making adaptable data editable and manageable within a localization process.

The National Centers were instructed on the principles and mechanics of translation/adaptation of PIAAC cognitive materials at the Lisbon NPM Meeting in October 2008, shortly before the release of the first cognitive materials (the literacy and numeracy link units). They received a general guidelines document prepared jointly by ETS and cApStAn, and attended an interactive training session on translation/adaptation procedures prepared jointly by DIPF and cApStAn. The training module included a detailed script, PowerPoint presentations, user manuals, various background and sample materials, and a hands-on session. It was shortly thereafter packaged and distributed as a "kit" so countries could replicate translation training locally.

Similarly to the general guidelines document for the BQ, its counterpart for cognitive materials stressed the need for very high quality translation in order to collect internationally comparable data – with the additional challenge, for cognitive materials, to "retain the *cognitive equivalence of tasks* as much as possible."

The general guidelines included the recommended procedure of double translation by two independent translators, followed by reconciliation by a third person. The team reconciliation approach (more suitable for questionnaires) was not advocated, but a review of the reconciled version by national domain experts was recommended as an additional quality-enhancing procedure.

[10] Portugal later dropped out of PIAAC.

[11] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

The general guidelines laid down requirements for translators and reconcilers, addressed security/confidentiality aspects, translation traps, the general principles for cultural adaptations (and detailed instructions for the adaptation of currency items), and explained the LQC processes that would follow the initial submission by National Centers of translated materials. It was explained that some PIAAC items have been taken from previous surveys (ALL and IALS) but with changes to accommodate the computer delivery mode. Countries were thus encouraged to use their existing national versions of these items as a basis for their translations, which would nevertheless be verified for equivalence to the PIAAC source version. It was also explained that countries would need to use specific software to enable the automated scoring of items using the “highlight in stimulus” response mode (Chapter 2).

Countries were also given item-specific translation guidelines (also referred to as “translation and adaptation rules” or “item-by-item notes”), conveniently echoed in the VFFs – the forms used to document the translation/adaptation and verification process. These guidelines were intended to draw the translators’ attention to possible terminology problems, translation traps, and issues for which adaptations were recommended, desirable, or ruled out.

At the NPM Meeting in Barcelona in March 2009, the National Centers were given workshops on the specificities of translating literacy units, numeracy units, problem-solving units, and reading components. The focus of these workshops was to familiarize translators with the guidelines for translating and adapting tasks. That is, in addition to stressing the importance of accurate translations, the workshops were used to emphasize the key role the construct plays in helping to develop the adaptation guidelines. In order to accomplish these goals, these workshops were used to provide a brief overview of the construct, demonstrate sets of specific items, and share and discuss specific guidelines for the proposed items

Throughout the localization process (from initial double translations to reconciliation, then post-verification review, layout adaptation and final check), the National Centers were assisted via the OTRS ticketing system. Queries were routed to cApStAn, ETS or DIPF as appropriate.

As for the BQ, countries had the opportunity after the Field Test to correct translation errors that had come to their attention. Furthermore, the Consortium provided each country with feedback based on the Field Test data that included recommendations to check certain specific items (where applicable). As for the BQ, modifications to the cognitive items at the initiative of countries were required to be restricted to those that were absolutely necessary, that is, to correct *errors*, and to avoid “cosmetic” changes, carrying the risk of negative impact on item functioning for items which had proved to work well in the Field Test.

4.4 International verification of the national versions – Field Test

4.4.1 Assignment specification, verifier training

The following was the key “mission statement” for successful localization taken from the PIAAC Translation and Adaptation Guidelines:

> In order to collect internationally comparable data in the study, the equivalence of all national versions is an essential requirement, which means that the translation of materials must be of extremely high quality in each of the national versions used by

participating countries. Within the assessment context, an additional goal is to retain the *cognitive equivalence* of tasks as much as possible, so that each item examines the same skills and invokes the same cognitive processes as the original version, while being culturally appropriate within the target country.

Essentially, verification is the LQC process put in place to check to what extent National Centers were successful in accomplishing the above objective, and correcting course as needed. Thus the verifiers' mission statement was to:

- Ensure linguistic correctness and cross-country equivalence of the different language versions of the PIAAC instruments
- Achieve the best possible balance between faithfulness to source and fluency in target
- Document interventions for both National Centers and the Consortium

The verifiers were selected from cApStAn's experienced team: They are native speakers of each of the target languages, highly proficient in English as source language and as working language to document their findings. They are trained to assess whether translation and adaptation guidelines are followed and to document possible deviations, insert corrections as needed and provide expert linguistic advice. They are knowledgeable about equivalence issues, translation traps and meaning shifts that are likely to affect response patterns in achievement tests. They also have experience in assessing the relevance of cultural adaptations in data collection instruments. They are all familiar with the use of "verifier intervention categories" and verifier comments in a standardized form.

Verifiers attended a two-day training seminar in Krakow, Poland, in April 2009, organized by cApStAn with the participation of DIPF staff (or the follow-up session organized in Brussels, Belgium, in May 2009). They were instructed about the PIAAC Item Management Portal, the OLT software, the particularities of the different instruments to be verified (BQ, literacy units, numeracy units, problem-solving units, reading components), the subtleties of verifying scoring definitions for the highlight in stimulus response mode and the numeric entry response mode. The training seminar included presentations and hands-on exercises.

4.4.2 Overview of verification procedures

The National Centers submitted reconciled XLIFF files (or Word files in the case of paper-based instruments) for verification via the Item Management Portal, together with the appropriate filled-in monitoring instruments (VFF for cognitive units, BQAS and Dynamic Text Rules Spreadsheet for the BQ, or DTRS).

Verifiers were instructed to compare each sentence of the target version of the instruments with the corresponding sentence in the English source version, and:

a) Examine whether the content of the items was equivalent across the two languages, with only appropriate and needed adaptations (for cognitive materials, this involved checking compliance with each item-specific guideline listed in the VFF).

b) Examine whether the target language was linguistically correct and struck the right balance between faithfulness to source and fluency in the target language.

c) When necessary, propose corrective action in the target language and document these interventions, in English, in the monitoring instrument. Documentation involved selecting an intervention category to identify the type of issue, selecting a severity code, and writing an explanatory comment (see below for details).

d) Verifiers also checked and intervened as needed on scoring definitions proposed by countries for the highlight and numeric entry response modes (in cognitive units) and on dynamic text issues (in the BQ).

e) Verifiers also checked national versions against the latest PIAAC errata list, maintained and regularly updated by cApStAn.

During the verification process, the need became apparent to refine the policy regarding the range of acceptable responses in numeracy items (for both the "exact match" and "number match" methods, see Chapter 5) In collaboration among the Numeracy Expert Group, DIPF and cApStAn, tables were prepared per country (or per group of countries sharing similar characteristics as regards, e.g., currency) in which the acceptable correct responses were listed for each item.

Likewise, during the verification process, a workflow was set up for error and exception management: corrupt file management, special requests by countries concerning units under verification or after final check, late submissions, upload of erroneous or incomplete files by countries, and so on. In hindsight, many of the problems were traced to the highlight response mode – a novelty in PIAAC. The presence of numerous and complex scoring definition "tags" in the XLIFFs made the files with highlight items more difficult to verify and subject to corruption. Furthermore, the workflow for scoring definition was not optimal, requiring too many steps: a) initial definition of scoring-related textblocks by country, b) then verification by cApStAn, c) then re-definition and re-verification in case of post-verification changes made by country and/or changes made at layout adaptation phase or at final check phase.

Verifiers were monitored and assisted by cApStAn staff, who also reviewed verified materials, liaising as needed with ETS/DIPF/ROA/CRP Henri Tudor on content and/or technical issues, before materials were "delivered" to countries.

Delivery took place via the Item Management Portal; countries were advised through OTRS when a batch of materials was verified, receiving precise instructions on how to further process the materials as well as a handy overview monitoring file.

These instructions are a convenient way to present the verification process in detail and are reused (in abridged form) in the two subsections that follow.

4.4.3 Detailed verification process – cognitive materials

Introduction – Process

The post-verification phase of the translation/adaptation/verification process for PIAAC assessment began after the verifier reviewed one or more batches of materials; the materials with suggested corrections and accompanying VFFs were made available on the IMP; and a Verification-Monitoring spreadsheet was provided, giving an overview of the verification outcomes for the verified batches. At this stage, it was the National Center’s responsibility to process the verification feedback and prepare for final check.

Background – Verification outcomes and how they were documented

PIAAC assessment materials were verified sentence by sentence, taking into account both general and item-specific translation and adaptation guidelines, with the aim to ensure the best possible balance between faithfulness to source version and fluency in the target version.

Verifiers’ suggested corrections were documented in VFFs, using a framework of intervention categories and severity codes (defining the nature and seriousness of identified issues). Figure 4.2 shows an example of a VFF showing a verifier’s intervention.

Figure 4.2: Example of VFF showing a verifier’s intervention

NPM COMMENT	VERIFIER INTERVENTION	SEVERITY CODE	VERIFIER COMMENT	DI
	Missing Info	2	'and places' missing from translation.	
	OK			

The severity codes have the following meaning:

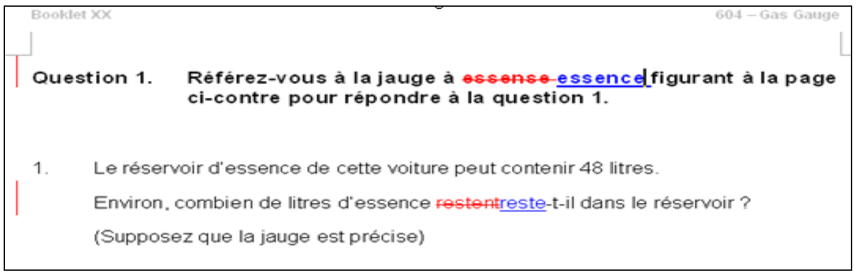
- **Code 1** - serious error (likely to affect item functioning – must be addressed – will be rechecked)
- **Code 2** - minor error (better to correct, but not crucial, so will not be rechecked).
- **Code 3** - suggestion for improvement (implementation left to the discretion of the National Reviewer).

Only Code 1 errors gave rise to follow-up at final check, as explained later.

The verifiers’ suggested corrections were mostly implemented in the materials. (Exceptions: in some cases verifiers reported layout issues that they could not correct, or made suggestions that were better not implemented but left to countries’ initiative. Such exceptions were always explicitly stated in the VFF: by default, verifiers' entries in the VFFs described problems that they went on to correct).

- Word units (paper-based) were corrected in "track changes" mode and needed to be processed by the National Reviewer (changes accepted or rejected or further modified). Figure 4.3 shows an example of a MS-Word file corrected in "track changes" mode.

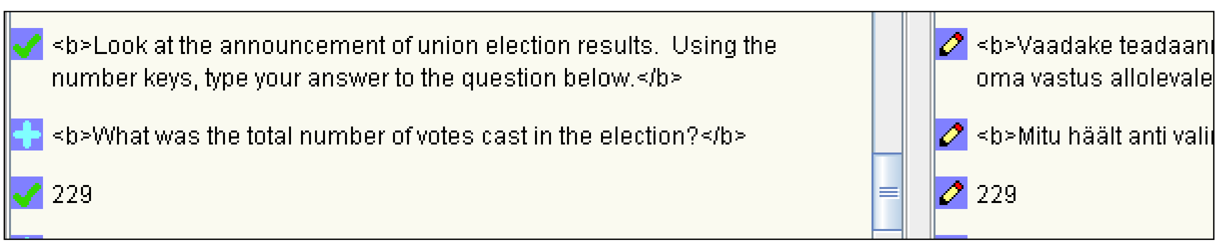
Figure 4.3: Example of Word file corrected in 'track changes'



- XLIFF files (computer-based) were verified using OLT, which does not offer the "track changes" mode. Instead, to show where verifiers intervened, text segments were marked (on the left side or "source" side) either "approved" (no changes made) or "translated" (some edits made).

- The National Reviewer did not need to take any action inside these files except if he or she disagreed with an edit. Figure 4.4 shows an example of a verified XLIFF file viewed in the OLT interface.

Figure 4.4: Example of verified XLIFF file viewed in OLT



Processing verification feedback – Step 1: Getting an overview

National Reviewers were advised to first consult the Verification-Monitoring spreadsheet, which provided a handy overview of verification outcomes for the verified batches. Figure 4.5 shows an example of a Verification-Monitoring spreadsheet, showing different verification outcomes for the units within a batch.

Figure 4.5: Example of Verification-Monitoring spreadsheet showing verification outcomes

PIAAC2009FT VERIFICATION LITERACY		COUNTRY: ZEDLAND	LANGUAGE: ZEDISH	CODE: zd-ZD
Unit	Name	XLIFF VERIFICATION Computer-based	LAYOUT ISSUES? Computer-based	NAT. I Comp FURTH
LINKING ITEMS (LITERACY)				
300	Employment Ad	DONE: MINOR EDITS ONLY		
301	SIGH	DONE: NO EDITS AT ALL		
302	Election Results	DONE: NO EDITS AT ALL		
303	Preschool Rules	DONE: NO EDITS AT ALL	YES	
304	Contact Employer	DONE: MAJOR EDIT(S)		
305	TMN AntiTheft	DONE: MAJOR EDIT(S)	YES	
306	CANCO	DONE: MINOR EDITS ONLY		
307	MEDCO Aspirin	DONE: MINOR EDITS ONLY		

- A column indicates for each unit whether it was verified with no edits at all, or with minor edits only, or with (also) major edits. This information was designed to save work: a unit verified with no edits at all did not even need to be opened – it was unchanged versus the version submitted for verification; a unit verified with minor edits only (severity code 2 and/or 3) was not further checked at final check.

- For all computer-based units, a column indicated the presence of residual layout issues (text that did not fit or display correctly, etc.). These were either noticed during translation and confirmed by the verifier, or noticed during verification. A "YES" in this column alerted Consortium technical staff that action was needed to fix a layout problem – no action by countries was required.

- For numeracy units and (a few) literacy units that include items with the "numeric response mode," an additional worksheet (Numeric-Entry-Scoring) indicated in column F the range of acceptable responses adapted to the country's situation (currency values, metric/imperial). This followed a Consortium decision to uniformly extend the range of acceptable "exact match" responses, taking into account variability in the way respondents "spell" numbers (dot or comma as decimal separator, comma or dot or space as thousands separator, dash to indicate "no cents"). Column G indicated whether this implied the need to implement additional correct responses; this was handled centrally by the Consortium – no action by countries was required.

Based on examination of the Verification-Monitoring spreadsheet, the National Reviewer could decide for which units to consult the VFF, which provided details on the corrections made. In turn, based on consultation of the VFF, the National Reviewer could decide which XLIFF files or Word files to open for processing.

Processing verification feedback – Step 2: Consulting/annotating the VFF

No post-verification entry by the National Reviewer was required in the VFF except in the case of disagreement with a Code 1 correction (Code 2 and Code 3 corrections could be freely accepted or disregarded/undone). In case of disagreement with a Code 1 correction, the National Reviewer was to enter in the "Discussion" column of the VFF a justification for not following the verifier's advice or correcting differently.

Processing verification feedback – Step 3a: Finalizing Word files

All Word files (paper-based units) with edits (including only minor edits) had to be opened so the corrections in track changes could be processed: accepted or rejected or (hopefully rarely) further modified. For reminder (see step 2), in case of rejection or further modification with regard to a Code 1 correction, the National Reviewer was to enter an explanation in the VFF.

Word files that were verified with no edits at all (as indicated in the monitoring spreadsheet) did not need to be opened – they were identical to the versions submitted for verification.

Finalized Word files needed to be uploaded to the IMP. This included Word files in which all corrections were just "accepted." There was no need to upload Word files that had been verified with no edits at all and which had not been further changed by the National Reviewer.

Processing verification feedback – Step 3b: Finalizing XLIFF files

The text in XLIFF files (computer-based units) was corrected during verification and the corrections/suggestions were not "provisional" (not in "track changes" mode). Verified XLIFF files only needed to be opened if the National Reviewer wished to undo or (hopefully rarely) further modify a correction – or to implement a suggestion listed in the VFF but not actually implemented by the verifier. As a reminder (see step 2), in case of undoing or further modification with regard to a Code 1 correction, the National Reviewer had to enter an explanation in the VFF.

Note that if the National Reviewer made post-verification changes to the stimulus text of a literacy unit that includes "highlight" items, this could affect the definition of textblocks for scoring. He or she had to send an OTRS ticket in that case.

Finalized XLIFF files were uploaded to the IMP.

Processing verification feedback – Step 4: Returning the annotated Verification-Monitoring spreadsheet

After the above steps are completed, the National Reviewer was to return the Verification-Monitoring spreadsheet with the "Further Edits" columns filled in. Figure 4.6 shows an example of a Verification-Monitoring spreadsheet returned by a National Centre, showing where the National Reviewer has made post-verification changes.

Figure 4.6: Example of Verification-Monitoring spreadsheet filled in by Nat. Reviewer, to show where post-verification changes were made in computer-based units

PIAAC2009FT VERIFICATION LINK-LITERACY		COUNTRY: AUSTRALIA	LANGUAGE: ENGLISH	CODE: en-AU	
Unit	Name	XLIFF VERIFICATION Computer-based (Outcome)	SCORING VERIFICATION Computer-based (Outcome)	LAYOUT ISSUES? Computer-based (DIPF/ETS will fix)	NAT. REVIEWER Computer-based FURTHER EDITS?
300	Employment Ad	DONE: NO EDITS AT ALL	DONE: NO EDITS AT ALL		
301	SIGH	DONE: NO EDITS AT ALL	NOT APPLICABLE		
302	Election Results	DONE: NO EDITS AT ALL	DONE: NO EDITS AT ALL		YES
303	Preschool Rules	DONE: NO EDITS AT ALL	DONE: MINOR EDITS ONLY		
304	Contact Employer	DONE: MINOR EDITS ONLY	DONE: NO EDITS AT ALL		
305	TMN AntiTheft	DONE: MINOR EDITS ONLY	DONE: NO EDITS AT ALL		
306	CANCO	DONE: NO EDITS AT ALL	DONE: NO EDITS AT ALL		

Final check

In the course of the final check procedure, units were reviewed in the following cases:

- Units were checked and corrected for residual layout issues and extension of acceptable "exact match" responses (technical final check).
- All units with major corrections (Code 1 corrections) were double-checked for correct implementation of such corrections (linguistic final check).
- In the case of computer-based units, this check was only needed for those Code 1 corrections for which the National Reviewer signaled disagreement in the VFF.
- In the case of paper-based units, this check was carried out on assembled booklets (PDF files), which were produced centrally by the Consortium.
- The workflow did not foresee another loop with units being returned to countries following the final check. Only the VFF was returned to countries upon completion of the final check, indicating for each Code 1 correction either "OK" or a comment suggesting that the issue was not satisfactorily solved. In the latter case, the National Reviewer still had the chance to address this, by making the recommended changes and re-uploading the affected units to the IMP.

4.4.4 Detailed verification process – BQ

Introduction – Process

The post-verification phase of the translation/adaptation verification process for the PIAAC BQ began after the verifier reviewed the nine XLIFF files of the BQ and annotated the BQAS; the materials with suggested corrections and accompanying BQAS were made available on the IMP and a Verification-Monitoring spreadsheet, was provided, which gave an overview of the verification outcomes.

At this stage, it was the reviewer's responsibility to process the verification feedback. There was no final check phase for the BQ.

BQ – Verification outcomes and how they are documented

The verifier compared each segment of the national target version (right-hand side of the XLIFF files) with the national source version (NSV, left-hand side of the XLIFF files). Both general and item-by-item guidelines were taken into account.

The verifiers' suggested corrections were documented in columns 16a and 16b of the BQAS, using the same framework of intervention categories as for the Direct Assessment, but without severity codes. Figure 4.7 shows an example of a BQAS with a verifier's intervention.

Figure 4.7: Example of BQAS showing a verifier's intervention

1 Int. Question No	2 International English Version	16a Verifier Intervention	16b Verifier Comment
B_Q15d	**Compared to your employer at the time, how useful do you think this training would be if you were working for a different employer? Would you say it was ...**	Missing info	Missing "Would you say it was ... .". Added by verifier.
	INTERVIEWER: Read categories to respondent.	OK	
	Much less useful	OK	
	Somewhat less useful	OK	
	Equally useful	OK	
	Somewhat more useful	OK	
	Much more useful	OK	

The verifiers' suggested corrections were mostly implemented in the materials. (Exceptions: in some cases verifiers made suggestions that were better not implemented but left to the country's initiative. Such exceptions were explicitly stated in the BQAS: by default, verifiers' entries in the BQAS described problems that they had corrected.)

- As for the Direct Assessment, XLIFF files were verified using OLT, which does not offer the "track changes" mode. Instead, to show where verifiers intervened, text segments were marked (on the left side or "source" side) either "approved" (no changes made) or "translated" (some edits made). The National Reviewer did not need to take any action inside these files except if he or she disagreed with an edit.

Processing verification feedback – Step 1: Getting an overview

National Reviewers were advised to first consult the BQ worksheet of the Verification-Monitoring spreadsheet, which provides a handy overview of verification outcomes. Figure 4.8 shows an example of a Verification-Monitoring spreadsheet with different verification outcomes for each of the BQ sections.

Figure 4.8: Example of Verification-Monitoring spreadsheet showing verification outcomes

PIAAC2009FT VERIFICATION BACKGROUND QUESTIONNAIRE		COUNTRY: ZEDLAND	LANGUAGE: ZEDISH	CODE: zd-ZD
Unit	Name	XLIFF VERIFICATION Computer-based	ADAPTATION ISSUES (CONTENT) / ROA ADVICE NEEDED	ADAPTATION ISSUES (DYNAMIC TEXT) / CRP ADVICE NEEDED
		(Outcome)	(ROA advice needed)	(CRP advice needed)
BQ	BQ Section AB (bq)	DONE, WITH EDITS	NO	NO
BQ	BQ Section C (bq)	DONE: NO EDITS AT ALL	NO	NO
BQ	BQ Section D (bq)	DONE, WITH EDITS	NO	NO
BQ	BQ Section E (bq)	DONE, WITH EDITS	YES	NO
BQ	BQ Section F(bq)	DONE, WITH EDITS	NO	NO
BQ	BQ Section G (bq)	DONE, WITH EDITS	NO	YES
BQ	BQ Section H (bq)	DONE, WITH EDITS	NO	NO
BQ	BQ Section I (bq)	DONE, WITH EDITS	NO	NO
BQ	BQ Section J (bq)	DONE: NO EDITS AT ALL	NO	NO

- A column indicated for each section of the BQ whether it was verified with or without edits. A unit verified with no edits did not need to be opened – it is unchanged versus the version submitted for verification (e.g. Sections C and J in Figure 4.8 above).

- Another column indicated the possible occurrence of residual adaptation issues that the verifier was unable to resolve and that (may have) required consultation with the BQ group (e.g. Section E in Figure 4.8 above). Usually an OTRS ticket was sent by cApStAn to the BQ group concerning such issues, and the issue was resolved or needed to be resolved between the National Reviewer and the BQ group.

- A last column indicated the possible occurrence of dynamic text issues, e.g., when the country commented in the DTRS that a given question did not require gender-related duplication or that a given past tense/present tense question required the introduction of an additional segment (e.g., Section G in Figure 4.8 above). Such issues were transmitted to CRP, and CRP contacted the reviewer concerning the best way to handle such issues.

Processing verification feedback – Step 2: Consulting/annotating the BQAS

The National Reviewer entered post-verification comments in the BQAS, for example, in the case of disagreement with a correction. If there was an additional iteration with the BQ group (see bullet points 2 and 3 of Step 1), there might be a Consortium comment in the BQAS. This would need to be taken into account when finalizing the BQ. In case of disagreement with a proposed correction, the reviewer sent the BQAS to ROA.

BQAS with post-verification comments were uploaded to the IMP.

Processing verification feedback – Step 3: Finalizing XLIFF files

The text in XLIFF files was corrected during verification and the corrections/suggestions were not "provisional" (not in "track changes" mode). Verified XLIFF files only needed to be opened if the National Reviewer wished to undo or (hopefully rarely) further modify a correction – or to implement a suggestion listed in the BQAS but not actually implemented by the verifier.

Finalized XLIFF files were uploaded to the IMP.

Processing verification feedback – Step 4: returning the annotated Verification-Monitoring spreadsheet

After the above steps were completed, the National Reviewer returned the Verification-Monitoring spreadsheet with the "Further Edits" column filled in. Figure 4.9 shows an example of a Verification-Monitoring spreadsheet returned by a National Centre, showing where the National Reviewer has made post-verification changes.

Figure 4.9: Example of Verification-Monitoring spreadsheet filled in by Nat. Reviewer, to show where post-verification changes have been made in the XLIFF files of BQ sections

PIAAC2009FT VERIFICATION BACKGROUND-QUESTIONNAIRE		COUNTRY: ZEDLAND	LANGUAGE: ZEDISH	CODE: zd-ZD	
Unit	Name	XLIFF VERIFICATION Computer-based (Outcome)	ADAPTATION ISSUES (CONTENT) / ROA ADVICE NEEDED (ROA advice needed)	ADAPTATION ISSUES (DYNAMIC TEXT) / CRP ADVICE NEEDED (CRP advice needed)	NAT. REVIEWER Computer-based FURTHER EDITS?
BQ	BQ Section AB (bq)	DONE, WITH EDITS	NO	NO	NO
BQ	BQ Section C (bq)	DONE: NO EDITS AT ALL	NO	NO	NO
BQ	BQ Section D (bq)	DONE, WITH EDITS	NO	NO	YES
BQ	BQ Section E (bq)	DONE, WITH EDITS	YES	NO	NO
BQ	BQ Section F(bq)	DONE, WITH EDITS	NO	NO	NO
BQ	BQ Section G (bq)	DONE, WITH EDITS	NO	YES	YES
BQ	BQ Section H (bq)	DONE, WITH EDITS	NO	NO	NO
BQ	BQ Section I (bq)	DONE, WITH EDITS	NO	NO	NO
BQ	BQ Section J (bq)	DONE: NO EDITS AT ALL	NO	NO	NO

Final check

Different from Direct Assessment units, there was no final check procedure for the BQ. The Verification-Monitoring spreadsheet and the BQAS with the National Reviewer's annotations were archived to keep a trace of the "history" of each national version of the BQ instrument, and reused when preparing the Main Study instrument.

4.5 International verification of the national versions – Main Study

The guiding principle of PIAAC Main Study translation/adaptation and verification activities was to control and limit the changes made by National Centers to their finalized Field Test national versions of assessment instruments, and carry out a verification focused on just these changes, with exceptions as needed and more extensive checks in identified "risky" cases, as well as a full verification of newly translated materials (the Main Study CAPI Workflow, Help Screens and Orientations).

The above scheme applied to "Phase I" of a two-phase process for revising and adapting the materials for the Main Study, devised in order to accommodate the tight timeline between the Field Test and Main Study. Phase I took place from May to November 2010, prior to analysis of the Field Test data, and focused on correcting issues associated with wording, scoring and layout that were identified by countries and the Consortium.

Phase II followed immediately after analysis of the Field Test data, from December 2010 to January 2011, and focused on identifying and correcting errors to the PIAAC instrumentation based on analysis of the Field Test data. During Phase II, a number of cognitive and BQ items that did not work well in the Field Test for a majority of countries were dropped. In addition, country review was allowed for a very limited set of cognitive items that functioned well for most, but not all, countries. Countries were asked to document the possible source of error and proposed solutions (beyond fixes which might already have been made during Phase I). A very limited number of last-minute changes were thus made at Phase II. These were discussed, approved and tracked, but not formally "verified" owing to the time pressure.

The rest of this section will describe the verification processes implemented during Phase I, where the great majority of Field Test to Main Study revisions were made. Countries were instructed on procedures at the NPM Meeting in Frankfurt in June 2010 as well as in the preparatory run-up to that meeting.

4.5.1 Main Study verification of literacy, numeracy and problem-solving units

The starting point of the process was the MMF (Main Study Translation-Adaptation-Verification Monitoring Form). One MMF for each of five batches (Link Literacy, Link Numeracy, New Literacy, New Numeracy, Problem Solving) was prepared and initially sent to National Centers with instructions to take note of the Main Study revisions and checks requested by the Consortium and add their requests for "national" changes (with a strong recommendation to limit these to corrections of errors, avoiding cosmetic or stylistic changes).

Note: specially customized MMFs were prepared for Hungary, which had dropped out of the Field Test process and rejoined for the Main Study (MMFs based on Slovak-Hungarian materials) and for the minority-language versions for Spain (MMFs with additional checks for Catalan, Galician, Basque, and Valencian materials which had not been final-checked at Field Test).

Countries' requests for changes were evaluated by the item development teams (sometimes with the assistance of verifiers who provided linguistic advice in a "pre-verification phase") and a Consortium recommendation for each request (approval, approval with caution, or rejection) was documented in the MMF.

On the basis of this preliminary work, the MMF was further filled out with cells to document and follow up on all "agreed revisions" (in both computer-based materials, in a first stage, and paper-based materials, later), using a color scheme to facilitate differential processing:

- Blue cells: text changes to be implemented by country and subject to verification and, if applicable, to final check.
- Yellow cells: layout changes or numeric scoring changes to be implemented by the Consortium's technical teams (DIPF or ETS) and subject to country's signoff.
- Mauve cells (in literacy units only): revision of text blocks for the scoring of highlight items, to be implemented by country and subject to a verification procedure (see later).

Figure 4.10 shows an example of a MMF documenting the Main Study verification process of cognitive units.

Figure 4.10: Example of MMF with blue and yellow cells showing the verification process

PIAAC MMF (MAIN STUDY TRANSLATION-ADAPTATION-VERIFICATION MONITORING FORM)
NEW-NUMERACY Language_Country: et-EE

Unit N°	Unit Name	Location	MS revisions and checks	Additional MS changes	Feedback on MS changes	AGREED CBA REVISION ACTION (WHO) + FOLLOW-UP	AGREED PBA REVISION ACTION (WHO) + FOLLOW-UP
634	Peanuts	Stimulus	NONE			NONE	N.A.
		Question 1 (C634P001)	NONE	Change "Mitu grammi (g) süsivesikuid sisaldab pakitäis pähkleid?" *Reason: better understandable*	FEEDBACK: You will be able to make this revisions to CBA using the tool to be introduced at the June meeting in Frankfurt. Those revisions will then be verified by cApStAn.	TEXT CHANGE - IMPLEMENTED BY COUNTRY, SUBJECT TO VERIFICATION Country comment: DONE Verifier comment: OK, no further changes needed Country post-verif comment/signoff: N/A Verifier final check: N/A	N.A.
		Question 1 - Scoring (Left Panel Numeric Entry Number Match)	NONE			NONE	N.A.
		Question 2 (C634P002)	NONE			NONE	N.A.
		Question 2 - Scoring (Left Panel Numeric Entry Number Match)	NONE			NONE	N.A.
635	Parking Map	Stimulus	REVISION CBa: The "3" representing the 3 km/mi mark should be aligned to the ruler.			LAYOUT CHANGE Implemented by ETS ETS comment: DONE Country comment/sign-off: It`s ok	N.A.

Countries were provided with a manual explaining how to process the MMF and with the technical instructions for accessing materials on the IMP (or on the online "Copernicus" in the case of problem-solving units) and checking changes made by the Consortium (yellow cells), making changes under their responsibility (blue cells), and, for the two literacy batches, checking/correcting the "text blocks" used for highlight scoring and running a full testing protocol on highlight items.

After this pass by countries, the materials moved to verification phase.

Verifiers had read-only access to the units (via the preview facility on the IMP or on Copernicus) and were instructed to check all blue cells in the MMF, making sure that agreed changes were implemented correctly (and not-approved changes were not implemented). Verifiers were further advised that the changes were approved or rejected by the Consortium based on information given by the country (not always very detailed or informative) and mostly with little or no knowledge of the language; therefore they were allowed to contradict or question the decision in cases where an agreed change could make the item easier or more difficult, linguistically poor, or causing an additional problem that was not taken into account by the country.

If the verifier detected no issue, the blue cell would be completed with markings that no further processing was needed (no need for final check). Otherwise, the verifier would describe the issue and suggest corrective action, and the blue cell would be marked for final check after post-verification review by the country.

Verifiers were also instructed to make use of the “Diff report” facility on the portal to detect and process any undocumented changes made by the country (this was a key feature to enable a “safe” focused verification procedure).

To verify the correct scoring of literacy items with the highlight response mode, a more efficient and focused procedure was put in place for the Main Study.

After countries revised their units, which included checking/correcting the text blocks and testing the highlight items, DIPF classified the national versions as low, medium or high risk, based on a review of the problems found at Field Test and of the quality and thoroughness of the Main Study scoring testing.

It was agreed that cApStAn would carry out a sample-based check of each country’s testing by performing a certain number of testing steps and checking that one received the same expected results. The list of testing steps to be performed was variable depending on the country’s classification. At minimum (low risk category): cApStAn tested all items for which the scoring rules were changed between Field Test and Main Study, any residual issues from Field Test testing (on a case by case basis), and three to five test cases chosen at random in other units than those already tested. For the medium risk category, cApStAn added six to nine test cases chosen at random in other units than those already tested. For the high risk category, cApStAn ran one or two additional test cases in each and every item.

Results were reported in the mauve cells of the Literacy MMFs with details in the separate scoring sheet where countries had documented their testing. For national versions that “passed” the validation procedure, countries were advised to nevertheless retest their scoring in case of text changes suggested by verifier that could affect the definition of text blocks (after implementing these at post-verification review stage). Such cases were clearly identified in the MMF. Countries that failed the validation procedure were asked to recheck and retest all highlight items and given further assistance.

4.5.2 Main Study verification of the BQ

As for literacy, numeracy and problem-solving units, in the Main Study, the principle was to verify only changes made to the BQ since the Field Test. The environment and process, however, were quite different. BQ sections were verified by reviewing and editing XLIFF files using OLT. The XLIFF files submitted by each country were specially prepared “partial” ones containing only the segments that countries needed or wished to change (not the entire BQ text), following a process of approval of national changes carried out with ROA.

When viewed in the OLT interface, verifiers see a “customized” (and approved) source version on the left and the country’s target version on the right. They were instructed to verify that the target texts are linguistically correct and match the customized source texts, make corrections as needed, and document these corrections.

The documentation and follow-up of verification corrections occurred in a new monitoring form created for the Main Study, replacing the unwieldy BQAS used in the Field Test. The “Main Study BQ Verification Report” form was designed to allow National Centers to easily identify where edits were made by verifiers and revert to their original translations. Follow-up columns were included for possible comments on verification issues by ROA (content issues) and/or CRP (technical issues, e.g., missing segments for dynamic text variants), who were invited to add comments after the verifier’s pass and verification review by cApStAn, and indicate or confirm any issues considered as crucial and thus subject to final check.

The edited XLIFF files for BQ sections and the BQ Verification Report form were then sent to country reviewers for post-verification processing. Countries were instructed that they could make post-verification edits (e.g., to undo or further modify a correction made by the verifier), that they were free to comment on their choices or not for non-crucial issues, e.g., minor linguistic defects, but were required to comment on issues marked as crucial and subject to final check. Countries were further advised that the most useful way to reply was, for example, “OK, we agree with verifier/ROA advice so no further change was made to the already corrected segment” or “We have changed to xxx because of reason yyy.”

After the country’s post-verification review, the files came back to cApStAn for final check. If an issue marked for final check was found not to be satisfactorily resolved, there could be one more iteration with the country before final signoff. Figure 4.11 shows an example of a BQ Verification Report with the documentation of a particular issue through all successive steps.

Figure 4.11: Example of BQ Verification Report with an issue documented through all successive steps

PIAAC MAIN STUDY VERIFICATION REPORT BQ			Language:	Country: Portugal	Code: pt_PT			
XLIFF Segment No	Identifier (XLIFF comment text)	Source version	Translated version	Verifier's suggested version	Verifier comment	CRP - ROA comment	Country post-verif comment /Signoff	Verifier final check
SECTION E								
8	Code: E_Q02a / Type: Question	In what kind of business, industry or service did you work? Please give a full description.	Em que tipo de empresa ou organismo trabalhava? Por favor, responda de forma detalhada. Por exemplo: ensino pré-escolar, tribunal, centro de saúde, câmara municipal, fiação de fibras de algodão, fabricação de tecidos de malha, preparação e conservação de peixe, fabricação de pão, comércio a retalho de vestuário, construção de estradas, etc.	Em que tipo de negócio, indústria ou serviço trabalhava? Por favor, responda de forma detalhada. Por exemplo: ensino pré-escolar, tribunal, centro de saúde, câmara municipal, fiação de fibras de algodão, fabricação de tecidos de malha, preparação e conservação de peixe, fabricação de pão, comércio a retalho de vestuário, construção de estradas, etc.	Consistency (according to the instructions that follow the question). Translation is much more specified than source version, includes examples of professions. Verifier did NOT change	ROA: Consistency: agree in principle, as long as this improves the clarity of the translated version Translation more specified than source: see above.	X We would like to keep the previous version to be consistent with other national surveys, like Labor Force Survey. This translation/extension was approved to the FT and would like to keep it.	OK, but then better be consistent using "empresa ou organismo" to translate "business, industry or service" in all occurrences --> We have thus changed to "empresa ou organismo" also in Section E segment 9 (see below) and in Section D segments 8 and 9 (see above) PT: Ok, thank you!

Note: for three national versions (Japan, Korea and the Russian Federation[12]), the countries requested and obtained approval to revise the entire BQ. The verification of these three versions was hence full rather than partial (changes only), but followed the same procedures described in this section.

[12] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

4.5.3 Main Study verification of the CAPI workflow, Help screens, and orientations

For the Field Test, these "ancillary materials" were translated centrally (by the Consortium) to mitigate the heavy translation workload for National Centers. For the Main Study, with the lesser translation workload, these translations followed the "decentralized" model and were thus subject to full verification.

The verification process was similar to the one implemented for the BQ and described in the section above: translation/adaptation by country in XLIFF files using OLT, verification via edits to the XLIFF files documented in a Verification Report form, post-verification review by country documented in the same form, but no final check in the case of these materials.

An important difference was an extra column in the Verification report form labeled "Special instructions, checks, errata." These were instructions for verifiers, specially prepared after thorough analysis of the files, with a view in particular to ensure key matches between elements appearing in these files with the translations used in test units (e.g., names of units, correct responses for "core" items scored by the interviewer, names of the problem-solving "environments" and tooltips, etc.), and in the CAPI interface. Figure 4.12 shows an excerpt from a Workflow-Help-Orientation verification report with some checks to be performed by the verifier.

Figure 4.12: Excerpt of WF-HELPS-OR Verification Report showing some checks to be performed by verifiers

PIAAC MAIN STUDY VERIFICATION REPORT
ORIENTATIONS

XLIFF Segment No	Source version	Special instructions, checks, errata
ORIENTATION - LITERACY		
22	Forward	To be understood as per context "Go forward", not as "Forward an email". The translation must match the term used in the country's PS Web environment
29, 30, 31, 32	Comma Period Slash for fractions Dash for negative numbers	These terms should be translated consistently across the Helpscreen-Literacy, Helpscreen-Numeracy, Orientation-Literacy and Orientation-Numeracy files
36	Thank you. Go to the next question OR Click here to go back and change your answer	The translation must match the translation used in e.g. 318-Civil engineering (New-Lit), Q1, when you click on any link.
40, 41, 42, 43, 57, 58	Email Web	"Email" and "Web" should match the terms used on the environment tabs in the country's New Literacy unit 327 Summer Streets. **NOTE:** you need to click on the link on the opening page to arrive in a screen in which both buttons are shown.
45	... To review how to answer a question, click on Help	Though formatting of references to buttons is not consistent in the source, make sure that it is in the target: either capitalization or other consistent formatting of NEXT, SPACE, BACKSPACE, HELP, BACK throughout the Helpscreen + Orientations files Note: source should be **To review how to answer a question, click on HELP.**

4.5.4 Main Study verification of paper-based materials

For the Main Study, paper-based materials were prepared and verified on an "easier" timeline than computer-based materials, from March to June 2011. Procedures differed for the various materials, which comprised:

- Paper-based test units, assembled in three booklets: Core, Literacy and Numeracy.
- Reading components exercises, assembled in one RC booklet
- Scoring Guides for the Core, Literacy and Numeracy booklets

For the paper-based test units, the Literacy and Numeracy MMFs used to document and approve Field Test to Main Study changes during Phase I were "exhumed" (as a reminder, these included changes to paper-based as well as computer-based units, which needed to be considered together) and complemented with new changes or checks resulting from the Phase II revision process. Countries were instructed to make changes under their responsibility (corresponding to blue cells in the MMFs, see earlier description of the process for computer-based units) to the final Main Study-Word files from the Field Test, in track changes mode, and send these to the Consortium's pre-press specialist, Danielle Baum. Danielle constructed initial PDF booklets based on the Main Study master versions, implementing the formatting/layout changes under Consortium responsibility (corresponding to yellow cells in the MMFs).

Verification of the correct implementation of agreed changes was carried out on these PDF booklets, with reporting of issues (and suggested corrections) in the MMFs. Verifiers also had access to the Main Study-Word files showing the Field Test>Main Study changes in track changes mode, which was handy if they needed to see the previous wording. In addition, verifiers could preview the computer-based version (where applicable) to ensure alignment of changes in paper-based with changes in computer-based. (Note: after a number of discrepancies were found in Spain-Galician and Spain-Basque materials between the PBA and the CBA, a full PBA to CBA identicalness check was carried out for these two versions.)

The MMFs were then used for post-verification review by countries, implementation of corrections by the Consortium's DTP specialist, final check by cApStAn on revised PDF booklets, and signoff by countries.

Note that the booklets also included a cover page and an introduction. These elements were also verified (classically, for equivalence to source and linguistic correctness), but the verification was documented and followed-up in a different MMF, together with the scoring guides and reading components, the "Guides and Booklets MMF."

The three scoring guides were verified in Main Study-Word, with suggested corrections implemented in track changes mode and followed-up (post-verification review by country and final check of crucial issues) via the "Guides and Booklets MMF." The "Guidelines for Scorers" section was verified classically, for equivalence to source and linguistic correctness. For the scoring sections, verifiers were instructed also to check that unit names, question stems and other elements matched the actual units, and that scoring instructions were properly adapted according to precise instructions inserted in the MMF. Figure 4.13 shows an example of a MMF with verifier's interventions in the scoring sections of a scoring guide.

Figure 4.13: Example of MMF showing verification interventions in scoring sections

PIAAC
MAIN STUDY 2011

Country: Portugal
Target language: Portuguese

PLEASE INSERT NEW LINES, IF NEEDED, TO DOCUMENT ADDITIONAL ISSUES

LOCATION	ENGLISH SOURCE	CONSORTIUM RECOMMENDATION	VERIFIER INTERVENTION	SEVERITY CODE	VERIFIER COMMENT
	Question 2: List two ways in which CIEM helps people who will lose their jobs because of a departmental reorganization.		Consistency	2	Name (in the singular form) not consistent with the booklet. Changed by verifier.
	1 Mentions BOTH of the following: • They act as a mediator for employees OR mediation • They assist with finding new positions [Note: Do not accept "Job Data Bank"; "Guidance"; "Courses"; or "Career Change Projects" These responses should receive a score of 7]	Attention: 'Job Data Bank', "Guidance", "Courses" and "Career Change Projects" (literal matches with stimulus - see the dot points in left column).	OK		
313 - International calls	**Question 3**: Identify the two situations in which you might have to dial 098.		OK		
	1 Mentions BOTH of the following: • For help connecting a call • (To make calls in countries where the list says) the service is manual AND/OR via operator 7 Any other response 0 Stimulus and response page(s) left completely blank **Correct answer**: 1	Attention: Maintain literal/quasi-literal matches with stimulus	Consistency	2	Translation of "call" not consistent with the booklet. Changed by verifier.

The reading components were treated as a special case, given that a subset of materials used in the Field Test was selected for the Main Study, with no Field Test-to-Main Study changes. The "verification" of these materials consisted of a careful check that national materials were correctly assembled (using the correct selected materials as in the Main Study international master version), from the verified and finalized Field Test versions. To ensure the latter, verifiers were instructed to randomly check one or two verifier interventions from the Field Test in each section (vocabulary, sentence processing and passage comprehension).

References

Dorer, B. 2011. Advance translation in the 5th round of the European Social Survey (ESS). *FORS working paper series,* (Paper No. 2011-4). Lausanne: FORS.

Harkness, J. A. 2003. Questionnaire translation. In J. A. Harkness, F. J. R. Van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 35-56). Hoboken, NJ: John Wiley & Sons.

Harkness, J. A., Villar, A. & Edwards, B. 2010. Translation, Adaptation, and Design. In J.A. Harkness, M. Braun, B. Edwards, T.P. Johnson, L. Lyberg, P. P. Mohler, B-E. Pennell, T. W. Smith (Eds.), Survey Methods in Multinational, Multiregional, and Multicultural Contexts (pp. 117-140). Hoboken, NJ: John Wiley & Sons.

Dept S., Ferrari, A. and Wäyrynen, L., 2010. Developments in Translation Verification Procedures in Three Multilingual Assessments: A Plea for an Integrated Translation and Adaptation Monitoring Tool. In J.A. Harkness, M. Braun, B. Edwards, T.P. Johnson, L. Lyberg, P. P. Mohler, B-E. Pennell, T. W. Smith (Eds.), *Survey Methods in Multinational, Multiregional, and Multicultural Contexts* (pp. 157-173). Hoboken, NJ: John Wiley & Sons.

Upsing B., Gissler G., Goldhammer F., Rölke H., Ferrari A., 2011. Localisation in International Large-Scale Assessments of Competencies: Challenges and Solutions. In R. Schäler (Ed.), *Localisation Focus, Vol.10- Issue 1*. Limerick, Ireland.

Chapter 5: Development of the Cognitive Items

Britta Upsing, Frank Goldhammer, Maya Schnitzler, Robert Baumann, Roland Johannes, Ingo Barkow and Heiko Rölke, DIPF; Isabelle Jars, Thibaud Latour, Patrick Plichart, Raynald Jadoul and Christopher Henry, CRP; and Mike Wagner, ETS

The implementation of the cognitive items for PIAAC faced several challenges. As stated before, PIAAC was the first international large-scale study to be conducted entirely on the computer. Therefore, existing link items from prior studies like IALS and ALL had to be converted from paper to computer. In addition, new items had to be developed both in literacy and numeracy to take advantage of the new possibilities of computer-based assessment. Further, an entirely new assessment domain, problem solving in technology-rich environments, was defined and items had to be developed. This was all done under a short timeframe in collaboration with participating countries that developed items on their own, as well as by combining item development teams from different countries.

To cope with these challenges, a multifaceted approach was taken, reusing existing item development and test delivery software to the extent possible and developing easy-to-use new software to fill in the gaps.

As a basis, the assessment software TAO was used. In the so-called electronic reading assessment (ERA) option of the PISA 2009 study, TAO was used with the Hypertext Builder, a graphical authoring tool for complex items. This approach was reused and extended for PIAAC, resulting in a completely new version called the CBA (Computer-Based Assessment) ItemBuilder. All in all, the following combination was used:

- Test definition, item sequencing, item questions: TAO
- Literacy linking items/stimuli: ItemBuilder
- New literacy items/stimuli: ItemBuilder
- Numeracy linking items/stimuli: ItemBuilder
- New numeracy items/stimuli: ItemBuilder
- Problem solving in technology-rich environments: TAO (new item type)

The software and the procedures to produce the items are described in more detail below.

5.1 Development of literacy and numeracy items

As outlined above, literacy and numeracy items have been produced using the CBA ItemBuilder software. For the linking items, the existing paper items were used as a draft, while the new items were built from scratch or drafted using other standard software.

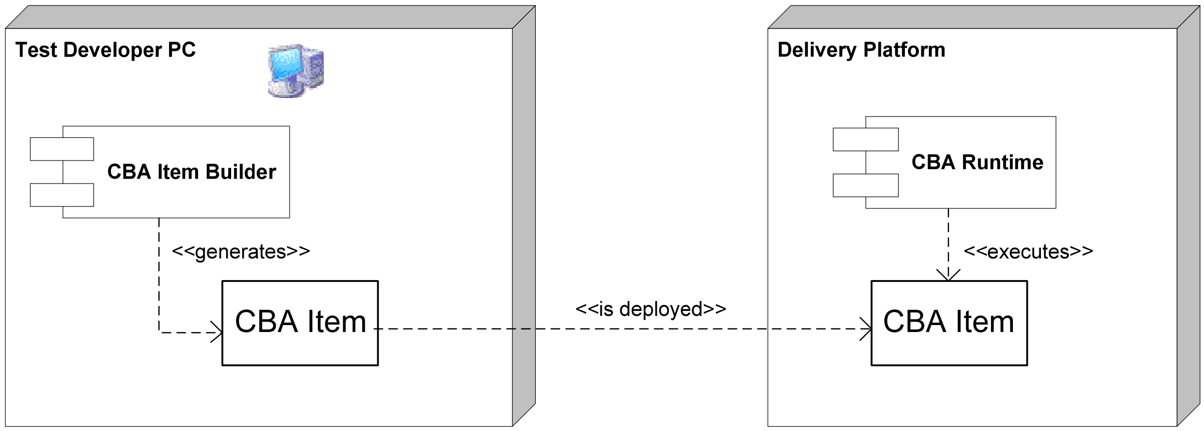
The goal was to produce an international *universal* item that could serve as a basis for country translations and adaptations without having to touch the layout. As it turned out later on, this often was not possible as there are great differences between languages in how much space they consume.

The CBA ItemBuilder is a graphical tool that enabled assessment domain experts to develop complex items in a what-you-see-is-what-you-get (WYSIWYG) manner without any programming. The CBA ItemBuilder consists of two major parts: the editor and an independent runtime environment. The editor allows for a WYSIWYG work style where you can drag item elements from a palette and freely drop them wherever necessary. Items designed in the editor are stored in an intermediate item description format. From this format, the executable item is generated using software generation principles. This ensures that the runtime for CBA ItemBuilder items can be changed without touching the editor in a relatively easy manner. The first versions of the runtime environment (used for PISA 2009 ERA) were based on Adobe Flash. This was changed in the preparation phase of the PIAAC study to a mix of standard Web technologies: HTML, JavaScript and SVG. The items produced with the CBA ItemBuilder could be used standalone as well as integrated into other assessment software. For simple assessment purposes, a graphical user interface was provided, allowing for editing and executing tests. In PIAAC, ItemBuilder items were used as stimuli integrated to TAO items. More information about the ItemBuilder software can be found in Rölke (2012). Here we only outline its possibilities.

5.1.1 The CBA ItemBuilder

The CBA ItemBuilder System consists of several interconnected components – see Figure 5.1 for an overview. The most prominent is the standalone CBA ItemBuilder. It is a desktop application based on Java and Eclipse technology. The CBA ItemBuilder was used to design and try out single items. Once an item was ready, a device-independent format was generated for later usage, the CBA Item. A CBA Item could be deployed on various platforms ranging from USB drives to the Internet.

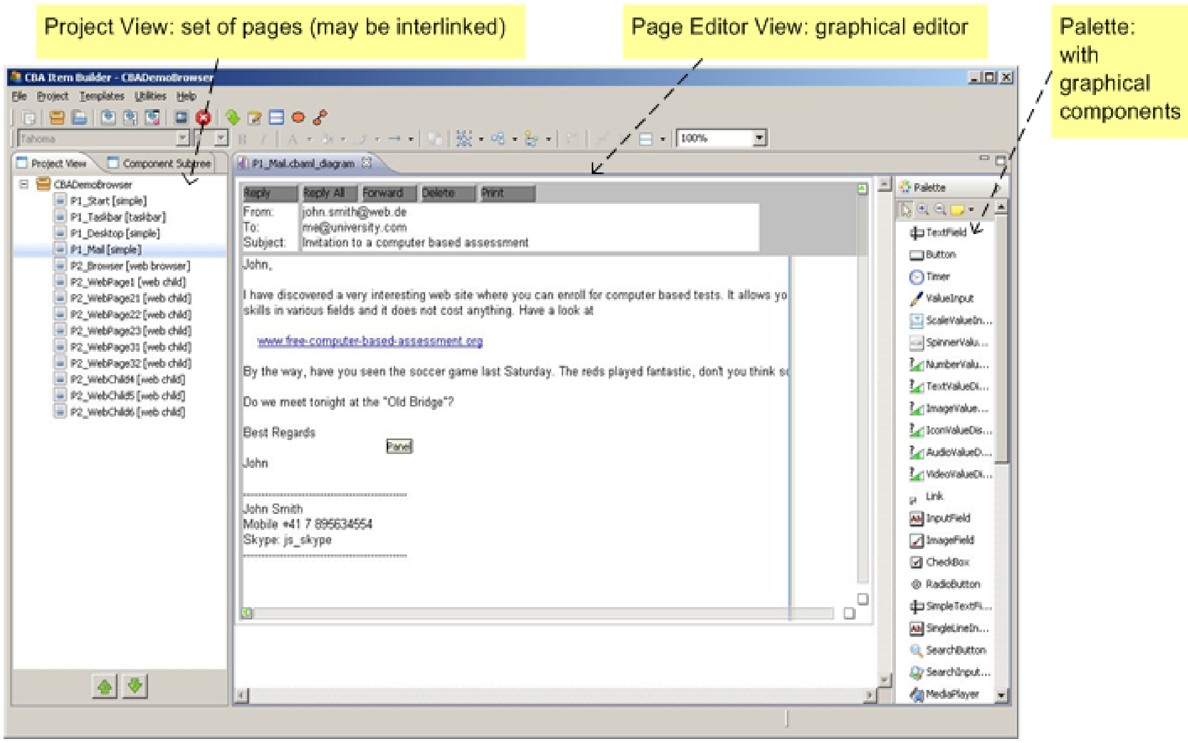
Figure 5.1: System architecture overview (M. Dorochevsky, Softcon)



CBA Items were delivered by a Java application server, usually JBoss. They could run on other servlet containers as well. As stated above, they consist of open Web standard: HTML code enriched by JavaScript and SVG. The Eclipse RAP framework (Eclipse Foundation 2012b) is used for graphical user interface components. Complex computations are being done on the server side, implemented in Java. CBA Items can be displayed on any current browser (e.g., Firefox, Chrome, Internet Explorer), provided JavaScript is enabled. For PIAAC, only Firefox was supported.

The CBA ItemBuilder offers a graphical user interface to compose stimuli and items via drag and drop. A stimulus could be used in one or several items, for example, in combination with different questions. As with other modern integrated development environments (IDE), it offered different *views* on the item at hand. Figure 5.2 gives an example for a mail client item in process.

Figure 5.2: CBA ItemBuilder graphical user interface



On the left side of the CBA ItemBuilder window, the *Project View* is shown. This view gives an overview on complex items that consist of more than one page. The Project View allows for a quick selection of all *Pages* that belong to one item (or project). In the middle of the screen, the *Page Editor View* is shown. This is the most important view. It allows for authoring of item pages or stimuli – to put it simply, something that can be shown on a computer screen later on in an assessment. The screenshot in Figure 5.2 shows the item as it is being edited. This is done in a mode that comes close to WYSIWYG but abstracts from images or design effects, for example. These can be included in a preview of the item.

On the right side of the image in Figure 5.2 is the *Palette View*. It contains all elements an item can be constructed with, for example, text fields, buttons or input fields. To apply such an element, the user simply drags it with the mouse and drops it in the Editor View. An example of an editing process can be found in the following subsection on so-called link items.

For an in-depth look at the CBA ItemBuilder, please refer to the mentioned literature (Rölke, 2012).

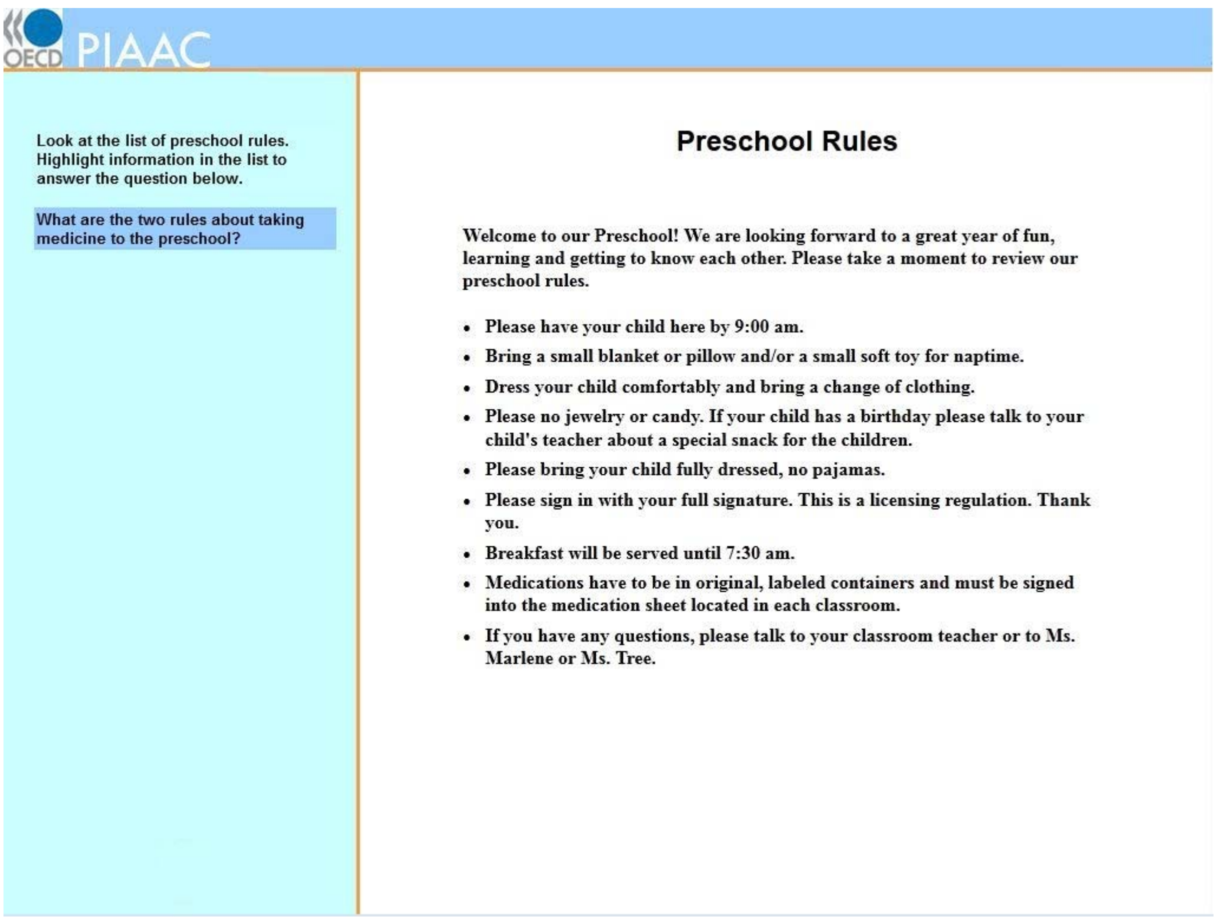
5.1.2 Link items

As the domains of literacy and numeracy had been measured in previous large-scale international surveys, it was a requirement that PIAAC link back to IALS and ALL. As a result, a set of linking items that had already been used in these studies was used in PIAAC. These studies have been performed as paper-and-pencil studies. The materials were therefore available as text documents in Microsoft Word format. The main challenge for the item development was to reproduce the paper item layout as closely as possible. Naturally, this could not always be achieved. To give just one example of the problems, the layout had to be changed from portrait-format paper to a landscape-format screen, and scrolling was not allowed.

Some materials like texts and images could be extracted from the existing paper items. Some images had to be redone, to add interactive areas, for example.

Figure 5.3 shows an example of a typical link item. It was originally derived from a newspaper article and was formatted for the paper assessment. For PIAAC it had to be reformatted to fit the screen layout and so on. To come as close as possible to the paper item that required marking as a means for answering the question, a new interaction mode called *multiple highlighting* was introduced. We come back to this interaction mode later on when dealing with the scoring of the items.

Figure 5.3: Link item example



5.1.3 New items

New items for literacy were developed by ETS in the US in cooperation with ACER in Australia. These items used features of the CBA ItemBuilder that were not utilized in the linking items. The new items in literacy focused on electronic texts, including Web pages, emails and discussion boards. This required the use of a simulated Web browser environment by the user. This environment was modeled in the CBA ItemBuilder. The ItemBuilder supported multiple stimulus “pages” within a single item. These pages could be linked via hyperlinks that were embedded in the item text. The runtime for these items supported maintenance of a hyperlink history, allowing the user to navigate back and forth among pages that had been visited.

The new items in literacy featured one unique response mode. Some items asked the user to click on a link in the text as his or her answer to the question. In these items, the scoring was based on the target of the hyperlink. Two pages were constructed, one for correct links and one for incorrect links. These pages looked identical to the test taker but had different identifiers internally. The test taker was given credit for a correct answer if he or she finished the item with the correct page showing in the browser.

New items for numeracy were developed by ETS in cooperation with members of the Numeracy Expert Group. These items had similar functionality and scoring mechanisms as the linking numeracy items. The only thing that distinguished these items from the linking items was the use of color images and artwork. The linking items, because of their legacy as paper-based items, were entirely in black and white.

5.2 Development of the automatic scoring software

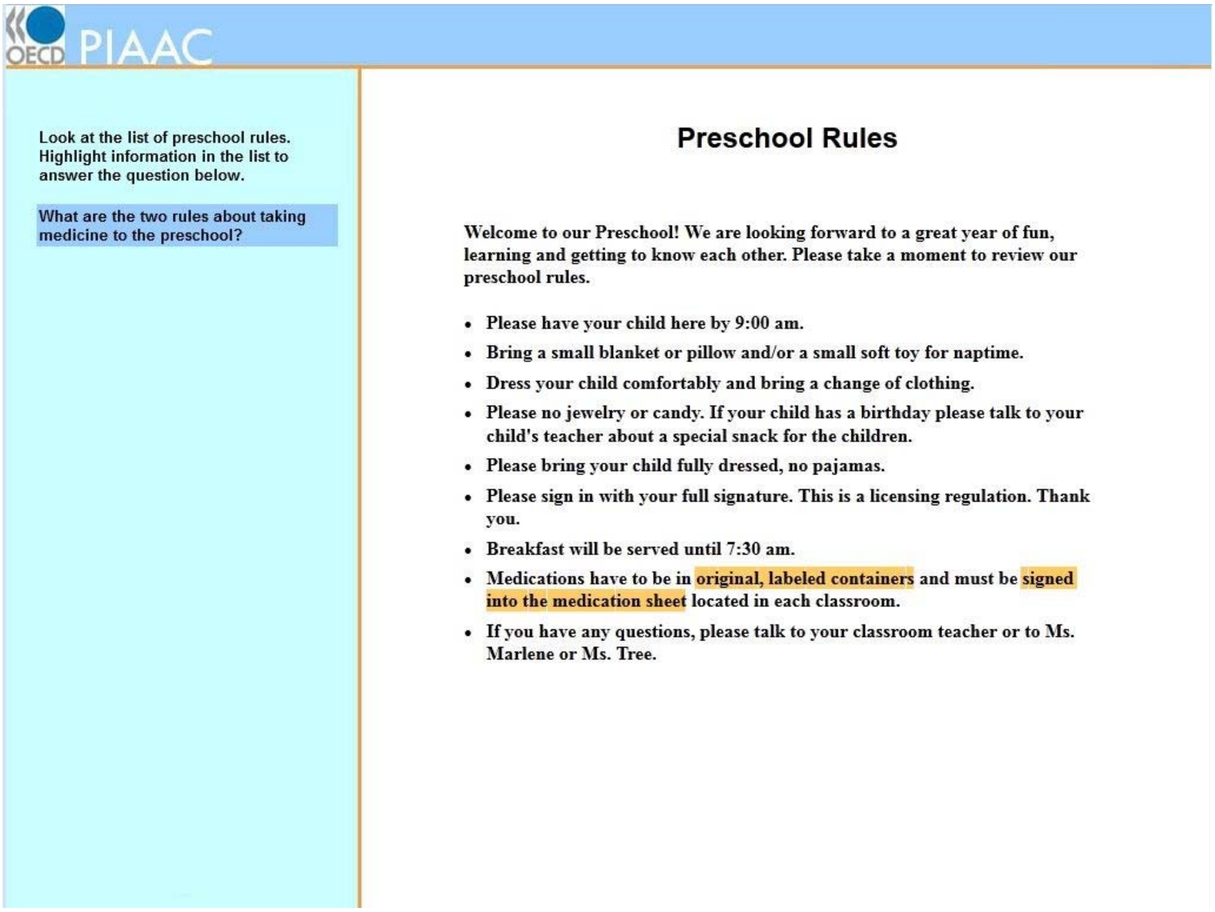
To enable adaptive testing in PIAAC for literacy and numeracy, those items had to be scored automatically and instantaneously by the platform. Various response modes were used which required developing different strategies for automatic scoring of test-taker responses. If a response mode included scoring rules with any textual or numerical information, they had to be adapted nationally. For those response formats, various workflows and (online) tools were developed to organize and support the national adaptations and the testing of adapted scoring rules.

5.2.1 Response modes in literacy and numeracy

Out of the PIAAC domains, literacy and numeracy included automatic scoring for adaptive testing and were based on a variety of response modes. Response modes could be divided into: i) those requiring interactions with the stimulus; and ii) those including interactions with the left panel of the PIAAC screen as shown below in Figure 5.4. The stimulus interactions were:

- Stimulus highlighting (items requiring the test taker to select a piece of text by clicking and dragging or by double-clicking a word)
- Stimulus clicking (meaning it is necessary to click on a graphical element, which usually becomes marked to indicate it has been selected; sometimes more than one element is clickable)
- Stimulus clicking link (used for new literacy items and refers to clicking on a link in a simulated Web browser environment)
- Stimulus multiple choice check box (clicking one or several check boxes that are included in a simulated Web browser environment used in new literacy items)

Figure 5.4: Scoring a highlighting item



The interactions on the left panel of the PIAAC screen were:

- Left panel single choice radio button/pulldown menu (used in new numeracy items and refers to clicking and selecting a single item in a group of radio buttons or in a pulldown menu, respectively)

- Left panel multiple choice check box (clicking one or several check boxes that are provided on the left panel)

- Left panel numeric entry number match/exact match (requires the test taker to enter number(s) into input box(es))

National adaptations were done mainly for: i) highlighting items because of the translation of textual information; and ii) numeric entry items because of adaptations to the national number format and/or currency system. Therefore, scoring testing efforts were focused on these item types. In contrast, clicking, multiple-choice and single-choice items were translated and adapted without affecting the scoring definition, that is, their scoring is the same as the one in the international version. In these cases, errors were assumed to be less probable because adaptation of scoring rule as one source of error is not relevant.

5.2.2 Automatic scoring of highlighting items

Stimulus highlighting items required the test taker to select one or more pieces of text by clicking and dragging or by double-clicking a word.

Highlighting items were usually scored by evaluating whether the minimum correct response had been selected, and, at the same time, whether the selection did not exceed the maximum correct response. Hence, each highlight item had a specified minimum and maximum response. The minimum response consisted of individual words or phrases that were identified as critical portions of a correct response. In general, the maximum correct response for highlight items included the entire line in which the correct answer was located, any or all of the line above the correct response, and any or all of the line below the correct response. However, the maximum correct response could not contain any incorrect information, as identified by test developers. So if any part of the line above or below had contradictory or incorrect information, it was excluded from the maximum correct response.

Highlighting responses were scored automatically by the system based on the definition of text blocks and the scoring rule referring to the text blocks. Text blocks define the parts of the text in the stimulus representing the minimum and maximum correct response. They were not visible for the test taker. Text blocks with the correct answer were already defined in the international item version. For the national versions of the stimulus, the text has been translated and the position and size of text blocks have been adapted by using a specific text block editor built into the CBA ItemBuilder software.

5.2.3 Automatic scoring of numeric entry items

Exact match items required the test taker to enter number(s) into input box(es) on the left panel of the PIAAC screen. Numeric entries were scored automatically by the system based on the definition of correct numeric response(s) included in the scoring rule. The exact match scoring method is equivalent to string match; that is, the system checked for character equivalence instead of numerical equivalence. For example, if a correct response for an exact match item was defined as ‘5’, an entry of ‘20/4’ would also have been scored as incorrect.

Number match items also required the test taker to enter number(s) into input box(es) on the left panel of the PIAAC screen. Numeric entries were scored automatically by the system based on the definition of correct numeric response(s) included in the scoring rule. The scoring method “number match” means that the response is correct as long as it represents the correct numerical value, regardless of the way the number is “spelled” by the test taker. For example, if a correct response for an exact match item was defined as ‘5’, an entry of ‘20/4’ would also have been scored as correct.

For some items, to retain realism, the magnitude of numbers and/or the number format were adapted for the national version. In this case, scoring rules also were adapted.

The Field Test scoring approach for number match items was revised to address several concerns expressed by experts and countries. In particular, the handling of decimal separators in the Field Test was considered to be too strict and unrealistic. Thus, the Main Study scoring approach introduced “double scoring,” which means that – within the system – the test taker’s response to an item was scored twice. So, before the system gave a final evaluation of the test taker’s answer, in all country versions it went through the following two scoring steps: i) The first scoring assumed a comma as decimal separator (i.e., acceptable thousands separators were blanks or periods), while ii) the second scoring assumed a period as decimal separator (i.e., acceptable thousands separators were blanks or commas). If at least one of the two scorings yielded a "correct," the response was considered correct.

Moreover, for the Main Study scoring, a so-called "strong mode" of the thousands separator(s) check was activated. This means that if the test taker used a thousands separator, the position of the separator needed to be correct to be considered a correct response. In general, only groupings by three digits were accepted. Groupings by four digits were acceptable only in the Japanese version (but not mixed groupings by four and three digits within a single number).

5.3 Scoring testing strategy

The automatic scoring procedures of the international versions of literacy and numeracy units were tested by the Consortium prior to distribution of the national versions. National language versions needed to be tested again by countries thoroughly, because for many items the definition of correct response(s) was adapted. This meant that scoring testing at the national level was especially important when the correct response included translated and/or adapted textual and numerical information.

The general rationale and procedure established for testing the international version were also the basis for testing the national version. They were revised iteratively during the international testing process. Basically, two sources of error were observed during international testing: i) errors at the level of item editing, that is, the scoring information was specified incorrectly by the item editor (specification error); and, ii) errors at the level of technology, that is, the software did not work accurately (implementation error). All detected errors were fixed, and the scoring of affected units was tested again successfully.

The testing was done manually, meaning the person responsible for testing completed a unit and item, respectively, as the test taker was supposed to do. Depending on the response mode, the tester used the keyboard for numeric entries or the mouse for selections to complete each item. For each response mode, a set of testing steps including the expected scoring result was defined to cover the most important test cases. When an item was tested, the tester gave a response as required by the testing step and compared the observed scoring result and the expected scoring result. Discrepancies between the observed scoring result and the expected scoring result needed to be documented and reported to the Consortium for debugging and for the Consortium to provide a revised version in the following testing iteration.

Countries were required to follow the international testing plan and for customizing the international test cases to their national versions as explained by the Consortium.

For the Main Study version of numeric entry items, implementing the adaptations and testing the adapted scoring rules was done centrally by the Consortium.

References

Rölke, H. (2012). *The ItemBuilder: A graphical authoring system for complex item development. in: world conference on E-Learning in corporate, government, healthcare, and higher education (ELEARN)*, pp. 344–353. Waynesville, N.C.: Association for the Advancement of Computers in Education.

Chapter 6: Development of Technical Support Tools

Britta Upsing, Frank Goldhammer, Maya Schnitzler, Robert Baumann, Roland Johannes, Ingo Barkow and Heiko Rölke, DIPF; Isabelle Jars, Thibaud Latour, Patrick Plichart, Raynald Jadoul and Christopher Henry, CRP; and Mike Wagner, ETS

6.1 Development of the Item Management Portal

The Item Management Portal was the central PIAAC portal for all aspects of item development, item management, translation, adaptation, scoring and layout testing, and so on. It consisted of several parts interconnected by several workflows. Workflows were available only to authorized roles (or users) of the Item Management Portal. Therefore, a user and rights management system was also available. (For example, a translation of an item could only be verified by an authorized verifier, and only after it had been released by the respective NPM.)

The Item Management Portal offered a multitude of different views to the central repository of all item-related information in PIAAC. It encapsulated a central database, a file server, a TAO Installation, and a CBA ItemBuilder server installation and it managed hundreds of gigabytes of data during the PIAAC study.

We cannot give a full overview of all aspects of the Item Management Portal here. Therefore we concentrate on specific parts to give an idea of the overall functionality.

Figure 6.1: Item Management Portal *Welcome Screen* for NPMs

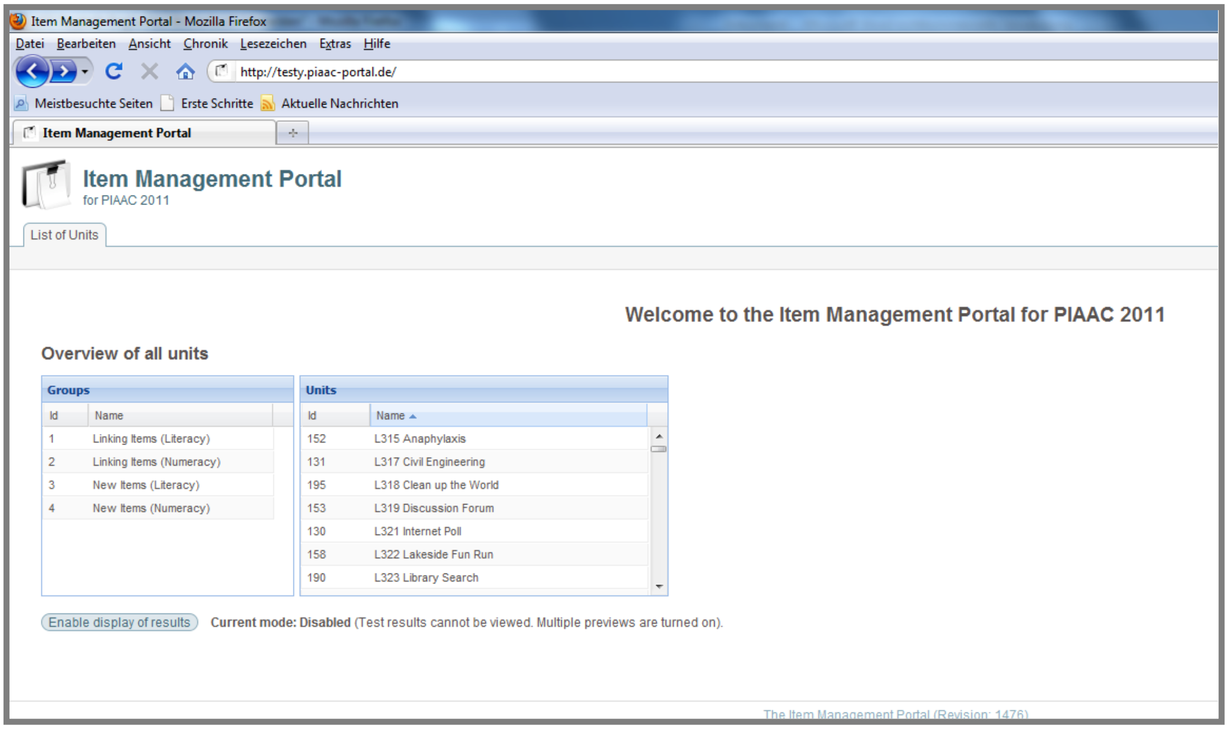


Figure 6.1 shows the so-called Welcome Screen of the Item Management Portal that was displayed after a NPM logged into the portal.

As the Item Management Portal was about items, several possibilities were at hand to access single items or for an overview of several items. For example, in the *Groups* box on the left, it was possible to select all link items in the field of literacy with one click. From those you could, for example, select a single item in the *Units* box on the right.

Figure 6.2: Item Management Portal with Item selected

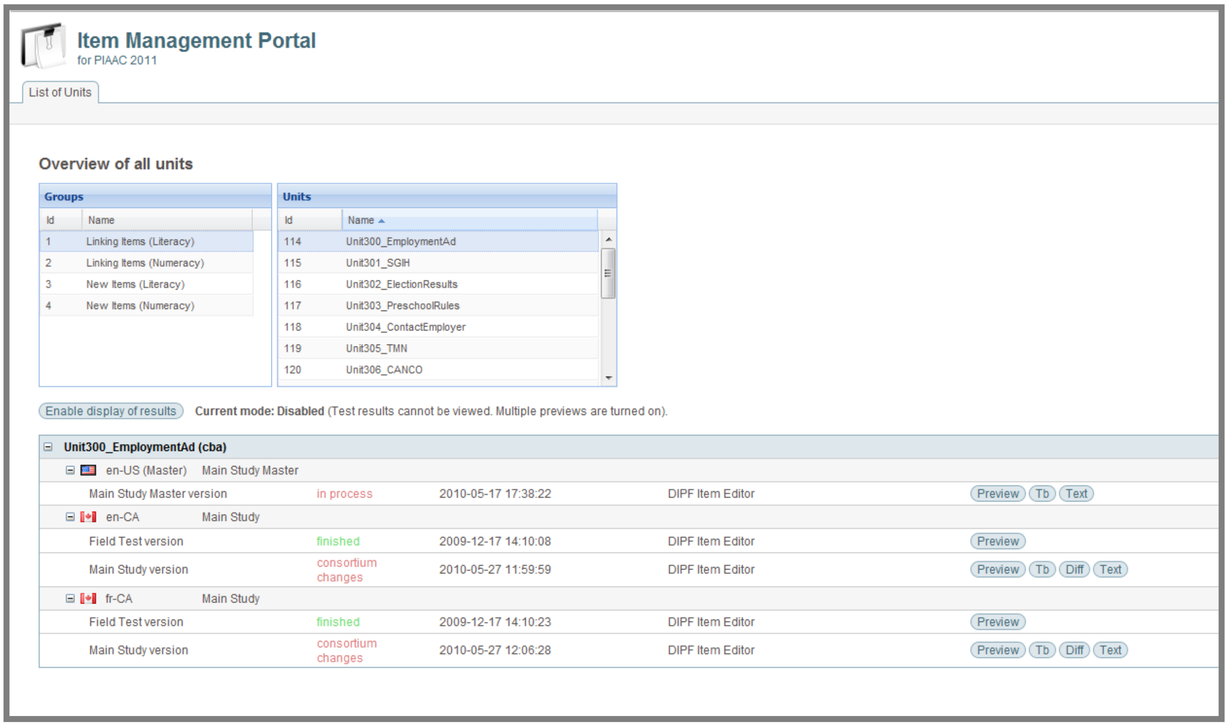
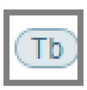


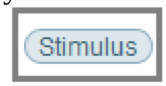
Figure 6.2 shows the Item Management Portal still in the NPM view, with one item selected. For this item, information about status and availability were displayed and several options were at hand depending on the step and the units' status in the overall process. The following list, for instance, shows the options that were relevant for preparing the main study version by revising the field test version:

- If you clicked on the  button, you would see the original version of the stimulus.
- If you clicked on the  Button, you would see a list of the text blocks included in the stimulus.
- If you clicked on the  Button, you would see all the changes made in the stimulus.
- If you clicked on the  Button, you would see a clean text version of the stimulus.

There were also several more buttons, not seen in Figure 6.2.

- If you clicked on the 

Button, you would see the discussion between your country and the Consortium.
- If you clicked on the 

Button, you would see the different inquiries that the stimulus consists of; you could choose the one you need to edit it.
- If you clicked on the 

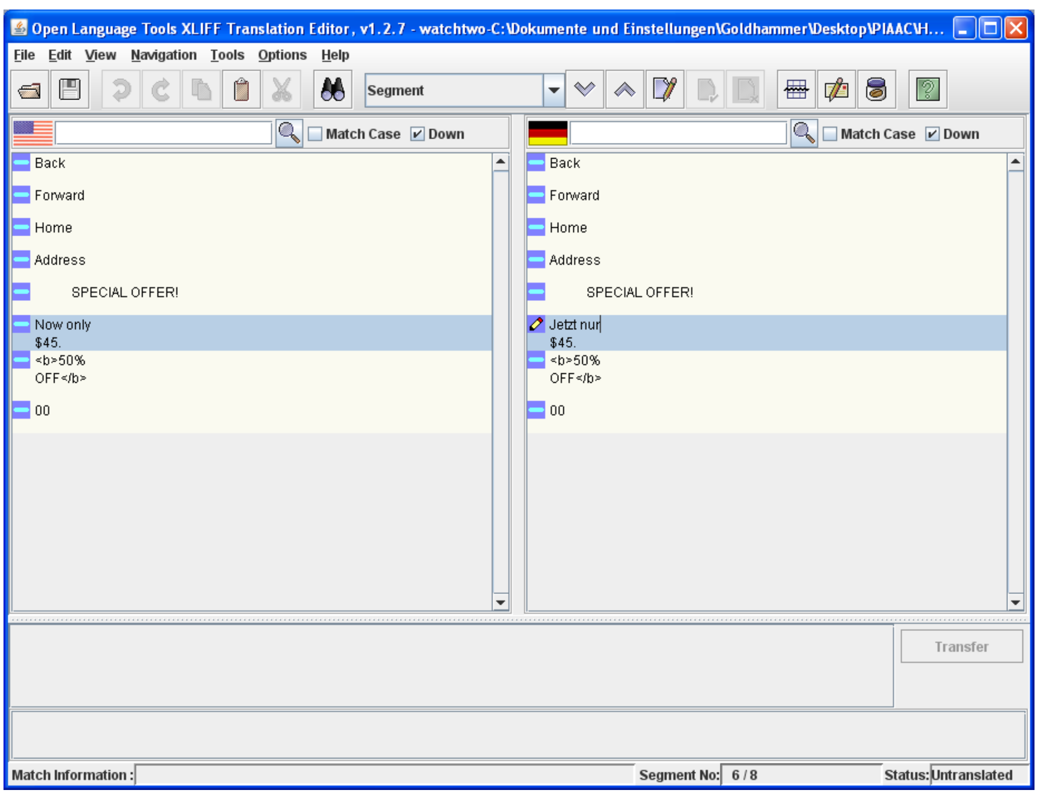
Button, you could download the stimulus.

Note that this is just a small part of the functionality of the Item Management Portal. Depending on your role and the status of the item development, other options were available such as upload and download of translation files, commenting, releasing items to the next step of the workflow and many more.

6.2 Translation and adaptation tools

For translation and adaptation of items, external tools were used rather than the Item Management Portal itself. This was to support offline work of these sometimes lengthy and time-consuming tasks without the need of a steady and reliable Internet connection. Two tools were provided: the Open Language Tool (OLT) and the Textblock Translation Editor (TBTE).

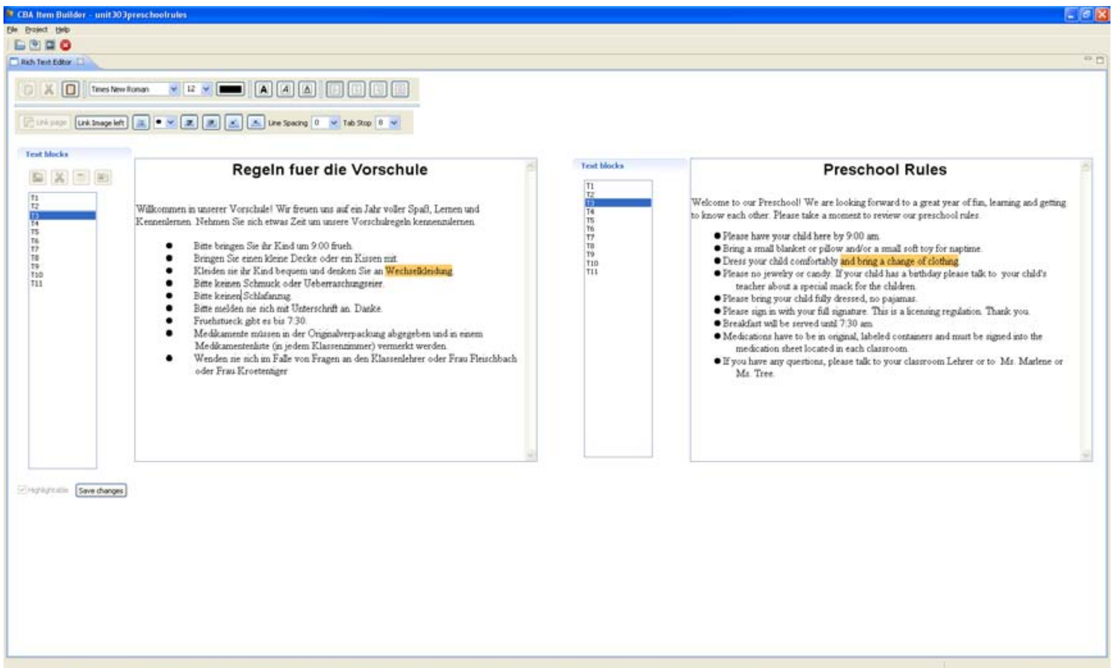
Figure 6.3: Open Language Tool



The OLT, shown in Figure 6.3, is an open-source translation editor implemented in Java originating from the company *Sun*. It had already been used for the PISA 2009 study and has been further developed since. The OLT builds upon the XML Language Interchange File Format (XLIFF), an XML standard for translations. The CBA ItemBuilder provides a built-in support for XLIFF. Doing the translation externally to the item authoring software rather than in the CBA ItemBuilder itself offered the advantage of limiting layout changes by translators to an absolute minimum. Further information about OLT can be found at http://open-language-tools.java.net/.

The purpose of the TBTE was to set the scoring definitions in translated highlighting items in accordance with the scoring definitions in the international item versions. The tool was to be used by the National Reviewer (NPM) in the context of setting the scoring definitions of a unit. In the workflow of the Item Management Portal, this step took place between *Reconciliation* and *Verification*. The TBTE essentially is a restricted version of the CBA ItemBuilder that does not allow for any change of layout and so on of an item other than changing text blocks. This is illustrated in Figure 6.4.

Figure 6.4: Textblock Translation Editor



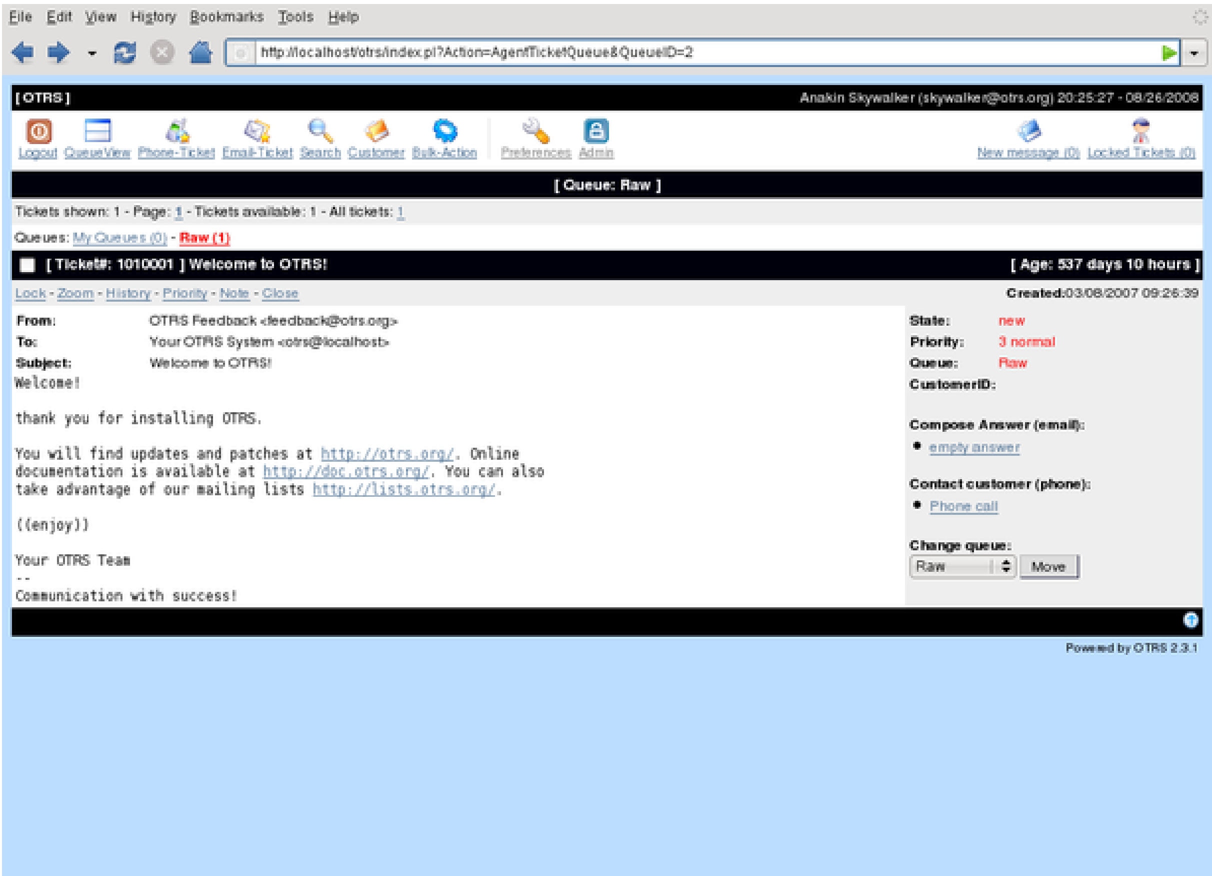
The TBTE was used to transfer the text blocks defining correct and incorrect parts of a highlighting interaction – see the chapter on scoring for details. The user could see both the original text and the translation and select one text block after the other. On the right side the original text block definition was shown. This was to be transferred to the translated text on the left side of the screen. Afterward the new text blocks became an integral part of the automatic scoring of the translated item.

6.3 Support tools

Support for users and stakeholders in PIAAC was mainly provided online, via mail and preferably via an instance of the ticketing portal OTRS (Open Ticket Request System). OTRS is an open source problem reporting system. More information is available at http://www.otrs.com.

Figure 6.5 shows an example view of the OTRS system. The image is taken from the OTRS user manual. The OTRS system could be used via any browser and/or via mail exchange. During the field phases, additional phone support was offered. The phone support was backed by the OTRS to make sure all support cases were stored in a single location for later reference.

Figure 6.5: Generic OTRS Ticket (from OTRS Manual)



Chapter 7: Development of the CAPI Questionnaire System

Thibaud Latour and Raynald Jadoul, CRP; and Mike Wagner, ETS

7.1 Introduction

This chapter describes the CAPI system that was used to implement the BQ of the PIAAC survey from a design and operational viewpoint. Because the CAPI system was built on TAO, and because TAO is intrinsically built around knowledge technologies, the general rationale of the BQ was for it to operate as a knowledge elicitation tool organized as an interview and implemented as a workflow (WF).

In knowledge technology terms, elicitation consists of populating a general model with everything one wants to know about particular topics, with specific data on particular observations. Technically, the information sought can be formalized as a knowledge model called an ontology. In a nutshell, an ontology formalizes a shared understanding of a domain by identifying a hierarchically ordered set of abstract concepts, the relationships tying those concepts together, and the properties that characterize them. In addition, concrete instances of the model are considered to describe particular individuals defined by the abstract part of the ontology. Giving a concrete example of a concept by creating a corresponding individual defines its nature. Specific values for each of the properties are assigned for concepts. The descriptions of these individuals are also part of the ontology.

In PIAAC, the CAPI system was conceived as a tool to describe these individuals as instances in an ontology corresponding to the BQ framework. Building such a tool involved creating the ontology with the knowledge domain corresponding to the BQ framework and properties of concepts associated to BQ variables; associating to properties some tools to collect the instantiation values (the questions and the questionnaire); organizing these tools in a consistent sequence (the interview process); and reporting the collected data. The following sections describe in details the specifications of the CAPI system to deliver the BQ in the household and how it was implemented into TAO.

7.2 CAPI system specifications and features

7.2.1 Overall description of the system

The overall goal of the BQ CAPI system was to support the data collection in households during the survey as well as support the creation of complex questionnaires. The system was designed for a series of different users such as system administrator, BQ authors and interviewers. To specify such a tool, and in order to capitalize on the experience gained in the ALL survey, a group was set up by the PIAAC Consortium gathering specialists from the Australian Bureau of Statistics,

Statistics Canada and Westat working in close collaboration with the teams at the Research Institute Henri Tudor, DIPF and ETS.

From a functional point of view, the CAPI system included a series of components supporting the authoring of the different elements of the questionnaire, that is, the questions and possible answers, as well as the specification of the question sequencing, along with translation in all the PIAAC languages and maintenance of successive versions. The system also supported the interview by presenting the questions to be read to the respondent by the interviewer according to the sequence defined by the BQ designers. Because countries had different constraints and processes regarding their national surveys and privacy regulations, the CAPI system enabled specific initialization decided at the level of the National Centers (NCs). The system also proposed a series of tools to help the interviewer and collect his or her remarks during the interview process.

The whole system supporting the PIAAC data collection on the field was run on laptops. As for other technological components, a series of technical and operational requirements were specified both for the implementation of the CAPI system and for the countries that had the responsibility to buy, set up, operate and maintain the hardware devices for their interviewers. Among these elements, security issues in terms of data confidentiality and integrity were scrutinized carefully. The usability, including the careful internationalization of the system for the interviewer, was also considered as they constitute an important factor that impacts the quality of the collected data.

Some countries were not equipped with a case management system suitable for PIAAC. To provide basic case management capabilities to these countries, basic functionalities were also envisioned.

The CAPI system was a standalone system that ran on a virtual machine (VM) within a laptop. To ensure that the countries were equipped with hardware that was sufficiently powerful with respect to the system, a questionnaire was elaborated for countries to describe the hardware they expected to use. In general, the system was not too demanding and modern laptops were sufficient, with a preference for fast CPU machine to accelerate the latency of the graphical user interface. Indeed, in order to control the overall duration of the interview, to keep the respondent focused, and to enable the interviewer to chain questions fluently, the Consortium fixed the maximum acceptable duration of the period between the validation of an entry and the display of the following question at two seconds. This latency included all intermediate response processing, that is, for branching or preparing precomputed responses.

The Linux Debian Open-Source Operating System (OS) was chosen for the VM and VMware, more particularly VMware Server – compatible with all standard host machines with OS Windows 2000, XP or Vista; MacOS; or Linux – was the selected VM technology. To keep the system as general as possible with respect to the various devices and softwares used by countries, the communication between the VM and the host system (on which countries were able to deploy their own third-party application such as case management systems) was restricted to file exchange and minimal script calls.

The internationalization and the management of languages and scripts were of utmost importance in PIAAC. Even if the CAPI screens were designed to trained interviewers solely, it nevertheless would strongly impact the quality of the interview and the collected data. All characters were encoded in UTF-8 (Universal Character Set Transformation Format – 8-bit). The preferred communication file format was XML, while CSV was also used for import and export, and ASCII text files were sometimes used to exchange small chunks of data during runtime. The translation system was based on exporting and importing an XML-based format called XLIFF (XML Localization Interchange File Format, see http://developers.sun.com/dev/gadc/technicalpublications/articles/xliff.html) that is compatible with the usual translator computer-assisted translation tools.

In order to ensure high quality and reliability standards for the collected data, security aspects attached to the CAPI system were thoroughly addressed both in terms of confidentiality integrity and accessibility. Data confidentiality was ensured by full encryption of all data and communication on the interviewer laptop as well as between the interviewer laptop and the consolidation points (at both national and international levels). The internal communications between the components of the VM and between the VM and external third-party applications that could run on the interviewer laptop depended on country requirements. To ensure data integrity and accessibility, a complete crash recovery system was set up. If the system crashed during the interview, it was possible to restart the questionnaire at the level of the last answered question and to provide information to the interviewer about the parts of the questionnaire that were already answered, the last answered question and the active language. The system also blocked most functionalities that did not pertain to the interview to avoid accidental termination of the CAPI program by the interviewer during the interview. However, if accidental termination happened anyway, the recovery system enabled resuming the system and the ongoing interview. Such a system was made possible using a fine-grained and very frequent input data auto-save capability. Auto-save was activated at the level of each input field every time the interviewer selected/entered input in a given question. Data integrity controls were also enabled by providing comprehensive log files of timestamped events and data at both question and flow levels for future external audits.

The CAPI system maintenance was ensured at three different levels: on the host system itself and the VMware software; on the hosted system and the associated software such as the Apache Web server, PHP and MySQL database; and at the application level, together with the configuration and collected data. The last two levels were maintained using versioned VM images and patches provided by the Consortium, while the first level was left under the responsibility of the countries.

Fluency of the interviews was also carefully addressed by providing to the interviewer a highly usable system with standardized and simple interface design, coupled with a comprehensive training of the interviewers. The data input mode was designed to be as easy as possible for interviewers. Indeed, the burden of using external devices such as mouses was eliminated by allowing the interviewer to input data, navigate in the question flow, and trigger functionalities exclusively with the keyboard. When answering questions with predefined responses, the interviewer was not forced to sequentially navigate through the list of options. Typing in the numeric code of the answer option on the keyboard automatically checked the corresponding value. Open questions with alphanumeric inputs worked the same way. Predefined standard answers such as refusal to answer or “don’t know” answers were made accessible using shortcut keys. In a similar fashion, triggering functions such as: pausing or leaving an interview, introducing comments and remarks, requesting help from the system or about a question, or navigation functions through the question flow (such as going backward and forward) in the questionnaire were also made by pressing shortcut keys on the keyboard. Along with some layout skins and question presentation conventions, the definition of these keys and shortcut mappings were made configurable for countries willing to customize them following their own interview standards. The overall design of the user interface was created with input from PIAAC participants in Australia, Canada and the United States.

In many situations, the respondent spoke more than one language with various fluency levels. Within the scope of PIAAC languages decided by countries, the system allowed the interviewer to switch to another questionnaire language that best suited the proficiency level of the respondent, independently of the system interface language used by the interviewer. Similarly, the interviewer had the ability to modify the interface language according to circumstances.

7.2.2 Running the questionnaire

The PIAAC CAPI system enabled the definition and execution of complex questionnaires that included conditional paths depending on previous answers provided by the respondent and a series of adaptive features that supported the interviewer during the whole interview process. Central to the design was the need to keep the interviewer focused on asking the question and collecting the answers from the respondent swiftly. To achieve this, strong support from the system had to be provided to avoid distracting the interviewer with unnecessary system manipulations, controlling the flow of questions, and ensuring response consistency by providing a series of automatic features. The BQ and the system were thus designed to include straightforward question types, navigation facilities, consistency checks, precalculated answers, adaptation of sentences to be read depending on the respondent qualities and previously collected information, and contextual information display and gathering functions.

Figure 7.1: Screenshot of a multiple-choice question illustrating the main questions components

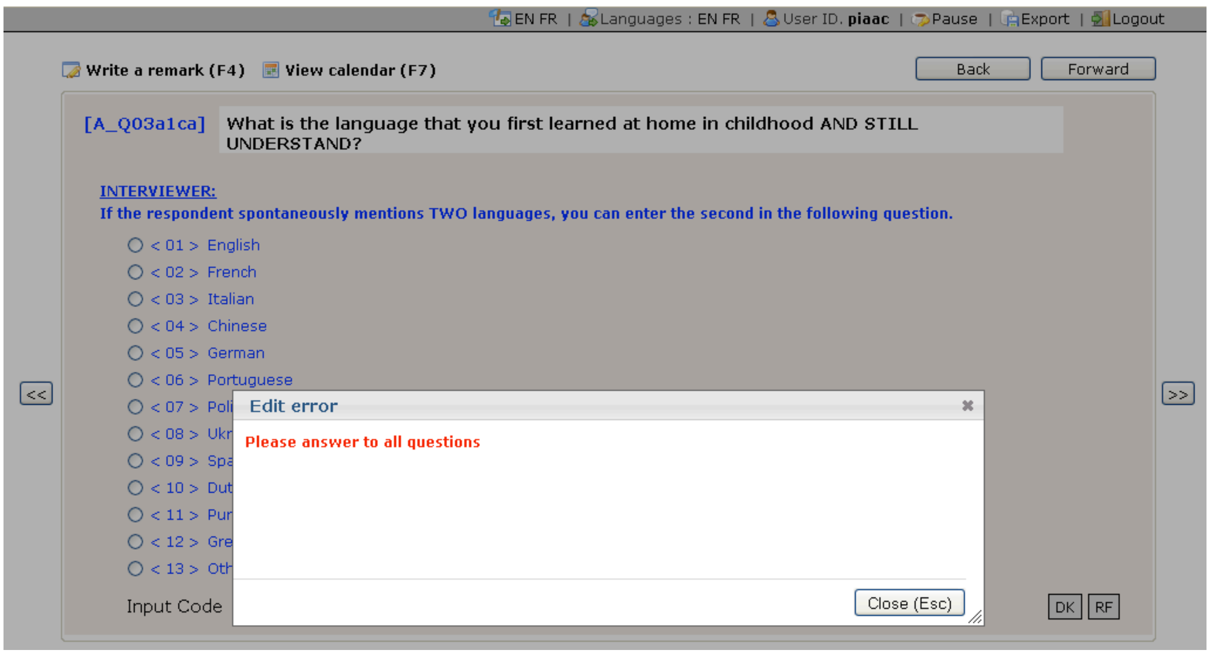
The BQ included several classical types of questions in terms of collected values, association with variables and in terms of layouts. Each question was associated to one or more variable. Depending on the nature of the variable, the associated value was a numeric response code, possibly associated to an alphabetic string or piece of text, or a numeric string. Whenever the range of possible value was closed and predetermined, a series of possible answers was provided along with their corresponding response code.

The system accepted both single-choice and multiple-choice questions, as illustrated in Figure 7.1. On the top, the question code was displayed along with the question texts to be read by the interviewer. Underneath the questions displayed instructions to the interviewer in a standardized color code (blue) indicating that the text should not be read. The series of options were provided with their response codes, among which the last one would be further specified. The globally predefined answers were displayed as the “DK” (don't know) and “RF” (refused) button associated with keyboard shortcuts. Navigation buttons were located on the left (back) and on the right (forward) of the screen. Response codes and navigation actions could also be entered using keyboard shortcuts.

Open questions collected textual or numerical answers. In many cases, free-text entries were used in conjunction with the fixed predefined answers to enable expanding the predefined list with new answers. In addition to the question-level coded answers, the system provided a series of globally defined answered to collect nonresponse for refusal or because the respondent did not know the answer. Responding to some questions was optional, but usually most questions were mandatory.

Figure 7.2 shows how mandatory questions were managed. Whenever more than one variable was associated to a question, various layout were available, among which was the table or array presentation. When meaningful, related questions scan were kept together in question blocks that were displayed together on a single screen. The visual layout was standardized with two basic display types indicating to the interviewer if the text should be read loudly as is to the respondent or it the text was an indication to the interviewer that it should not be read to the respondent.

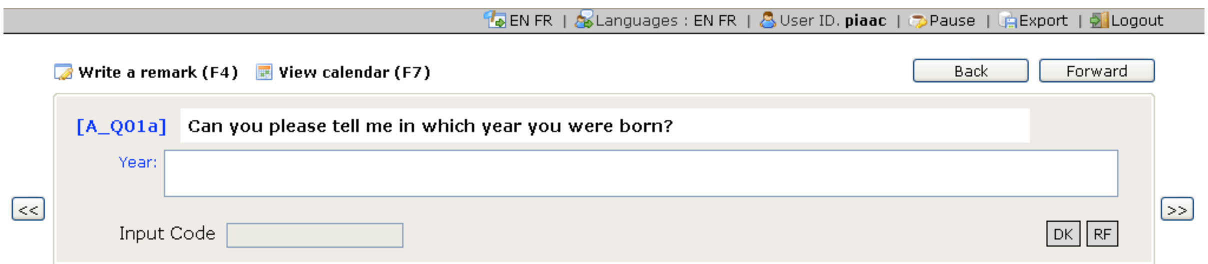
Figure 7.2: Screenshot of a mandatory question



The question sequencing in the PIAAC BQ followed a complex set of branching rules based on variables, either collected or calculated during the interview, loaded at initialization, or set up as global constants. Branching rules consisted of logical expressions that triggered a jump to a target question when a condition specified as a logical expression on any single variable or a combination of them was evaluated as true. Contrary to some other well-established CAPI systems, the one built on TAO for PIAAC did not evaluate the rule prior to displaying a question as a display condition or a skip rule. In the PIAAC system, the branching rules were always evaluated after the question was answered to determine what question to display next. In the absence of routing definition, the system assumed a linear sequencing and proceeded to the next question in the questionnaire definition order. Depending on country adaptations, the question flow and routing definition differed substantially from the flow of the international master BQ.

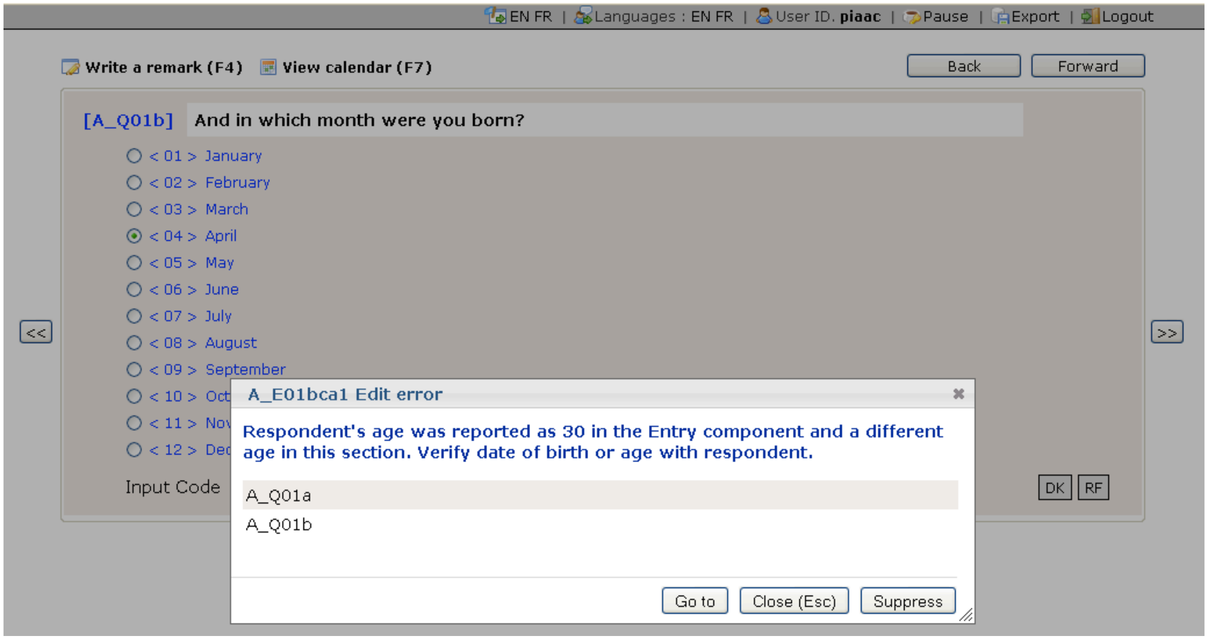
In case the respondent made a mistake or wanted to revise one of his or her previous answers, or when a consistency check forced the revision of a previous question, the system enabled going backward in the interview to modify the response. Such operation was tracked for later quality audit, and, when proceeding forward again, the routing was evaluated again with respect to the new data. Whenever the new question path differed from the previously explored one, all impacted values were invalidated but kept in the database.

Figure 7.3: Screenshot of a numeric entry to which is associated a range constraint and a consistency check



In addition to the internal consistency of variable inputs that were defined as constraints (such as numeric ranges or string length – for instance, when asking the birth year of the respondent as illustrated in Figure 7.3 – where the numeric value must be in a certain range), the consistency of responses with respect to context variables or previously collected data was maintained throughout the interview to avoid contradictions. These checks could be defined at the level of each question if needed. They were defined with a logical expression based on variable values and state, which triggered, if evaluated as true, the display of a piece of text explaining the nature of the problem and the set of previous questions that contradicted the current answer. Figure 7.4 illustrates the detection of a consistency violation concerning the age of the respondent. The BQ definition enabled specifying if corrections were mandatory or not. In the former case, the interview could not go further without eliminating the contradiction in either the previous or current responses. Similar to branching rules enabling the calculation of question routing, the consistency checks were evaluated after the question was answered before proceeding to the next question.

Figure 7.4 Screenshot of an edit screen appearing after violation of a consistency check



In some cases, the value of variables could be inferred from previously collected data. These situations were defined using inference rules, or auto-filling rules. Such rules were built on a logical expression based on variables that triggered the assignment of another variable if the expression were evaluated as true. The calculated value was obtained using a combination of string and arithmetic operators on other variable or constant values. Because inference rules were independent of the routing, auto-filled variables were sometimes needed irrespective of the previous step in the flow. Therefore, such rules could be evaluated before of after the display of the question to which they were associated. The value of the inferred variables could be used subsequently in any routing, consistency check or adaptive text.

Adaptive texts were specified as dynamic text rules that enabled substituting strings or substrings using variables or conditions on preexisting data. Adaptive texts were used to display the texts for the interviewer in the exact form they were to be read, taking into account the precise context of the interview. They were intended to prevent the interviewer from making on-the-fly adaptations of the discourse and maintaining good fluency of the interview. Typically, the adaptation concerned temporal context, taking into account the situation of the respondent (displaying “When you were at school” if the respondent left school, instead of “At school” if the respondent was still a student), or took care of the gender and, in some cultures, of the polite address depending on the respondent’s age.

Dynamic text provided a specific format in order to specify which part of the text had to be replaced or not, following a set of conditions. For instance, if the gender was coded as female, some parts of the question text had to be adapted accordingly. In the BQ, and in order to maintain the clarity of the text, the definition of the rule was independent from the string substitution. In the PIAAC system, only a variable was included in the text constant as a placeholder for the substitution. The content of the variable was obtained using an inference rule that was typically evaluated a priori, that is, before displaying the question. While all other rules were only dependent on country adaptation and related to the structure of the questionnaire, dynamic texts modified the content of the questions and were also language-dependent.

Such dependency had several consequences: The localization of the placeholder on the constant part of the text could vary from one language to another (according to the structure of the sentence), and some dynamic adaptations could require a different number of variables in different languages (due to conjugation for example). In principle, for a given piece of text to be translated, the definition of the required variables, their localization in the text, and the specification of the rule condition was part of the translation and should have been made by the translator. However, most translators were unfamiliar with the formal aspects of inference rules and string substitution. An easier solution was put in place where the entire piece of text was changed according to the rule for a predefined set of situations. For each situation (defined in the form of a hidden predefined rule), the translators were asked to proposed complete texts for all variations. The translated files were then post-processed to generate the correct inference rule and substitutions in the BQ definition.

Besides the content of questions, predefined answers and response codes, or rules, the system also enabled authors to define instructions related to the questions. Such instructions were located on the screen according to the reading flow of the interviewer – before the question text, after the question text, and after the response area. The objective was to facilitate the interviewer reading down the screen and getting the instructions in order related to what he or the respondent had to do. For example, an instruction directing use of a show card was given before the question text. This prompted the interviewer to supply the respondent with the show card before proceeding to reading the question. The most common instructions, directing the interviewer on what to do to complete the question, came after the question text. For example, a special coding or probing instruction may have fit in that location. Having read the question, the interviewer received instructions on how to handle input or question follow-up in position with its relevance. Instructions after the responses were special to guide interviewers on what to do next. Rich text was used as a way to control the visual presentation of the instructions for the interviewers.

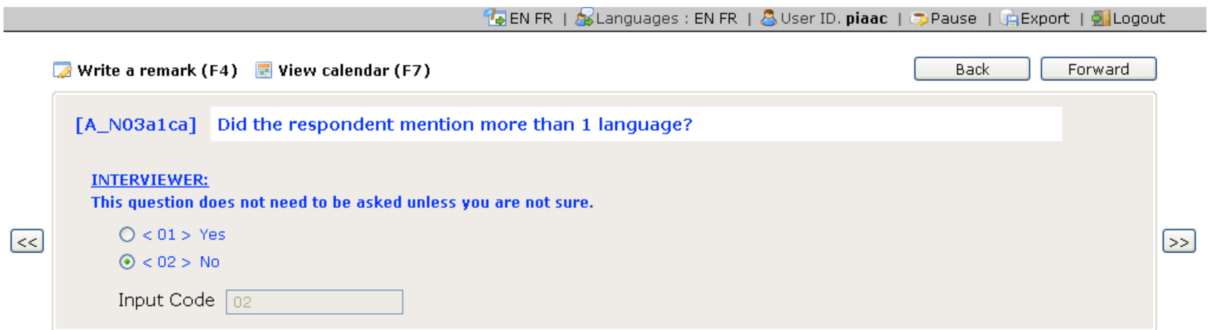
In addition to on-screen instructions, the system also enabled BQ authors to create supplementary helps and instructions that could be consulted optionally by the interviewers. Such helps usually consisted of more precise definitions of some concepts mentioned in the questions that the interviewer could use to give precisions to the respondent. The question-related help material was made visible in a modal window that popped up when the user requested it.

7.2.3 Running the interview

The CAPI system could be initialized by importing a set of predefined variables, for example, personal data like name, address, and so on of the respondent. Depending on the countries, these data were passed by the basic case management system provided with the PIAAC system, or generated from a third-party more sophisticated case management system installed on the laptop. When the questionnaire was executed, variables corresponding to initialization data were prefilled and usable in rules expressions or visible as prefilled responses to questions. Upon country adaptation, the prefilled responses might then be verified with the respondent and corrected as the case might be. In case of a breakoff, the questionnaire was populated with responses from the previous session when the interview was resumed at some later time.

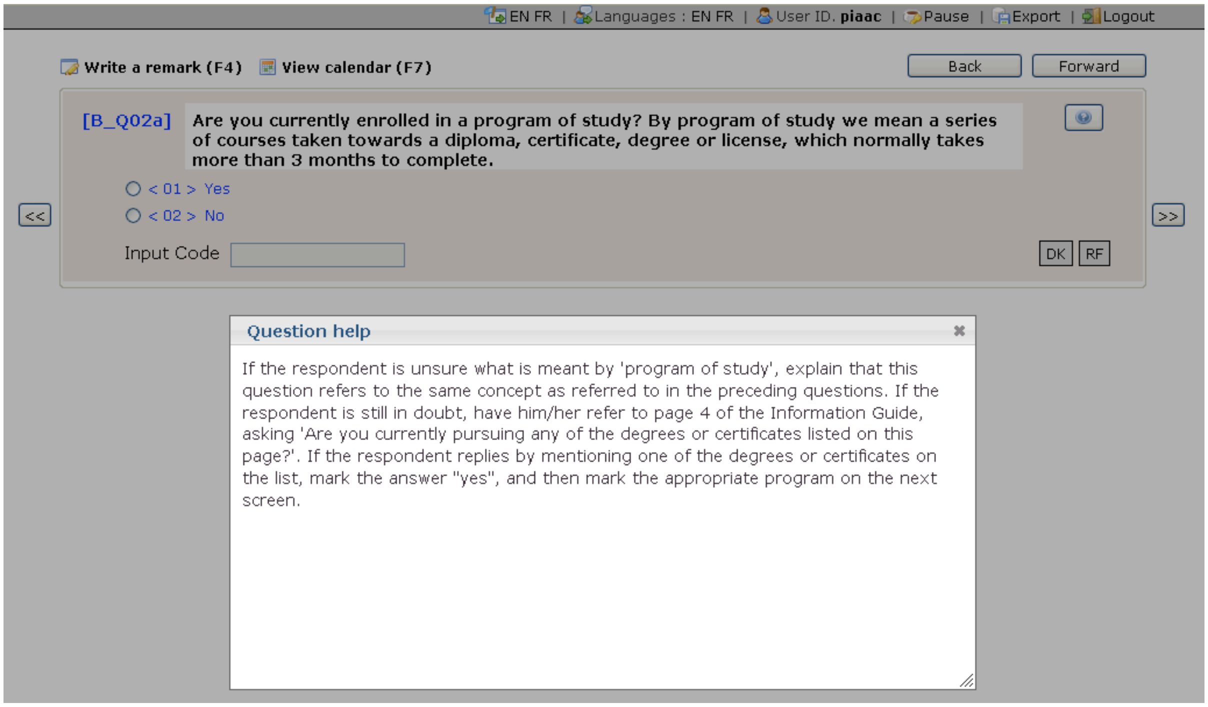
During the interview, the questions were displayed according to the sequence defined in the BQ. A priori inference rules were first calculated, if any, and the new variable calculated. Then the text substitutions were executed to compose the possible dynamic texts that might appear in any string on the screen, be it question, instruction, help, predefined answers, and so on. The substitution was also performed each time the interviewer toggled between languages or every time a question was revisited during navigation. The later situation could arise if a previous answer was changed that modified the context of adaptation of a subsequent adaptive text of a question the interviewer was coming back to again. The inference rule could also prefill the answer to the question that the interviewer might change.

Figure 7.5: Screenshot of a checkpoint question that is not read to the respondent but contains only interviewer instructions



Once displaying the question, the system showed instructions to the interviewer with a specific, regular visual presentation enabling the distinction between different types of instructions and questions read to the respondent. Answers might sometimes be read aloud to the respondent, depending on similar rendering rules as for questions. As illustrated in Figure 7.5, instructions could also be part of special questions that were not read to the respondent but served as checkpoints during the interview. The system also provided help buttons to give access to help information attached to the system, the interview, the current question and the currently used interviewer interface at any moment during the interview. Figure 7.6 shows an example of help material provided to the interviewer.

Figure 7.6: Screenshot of a model screen displaying help material to the interviewer



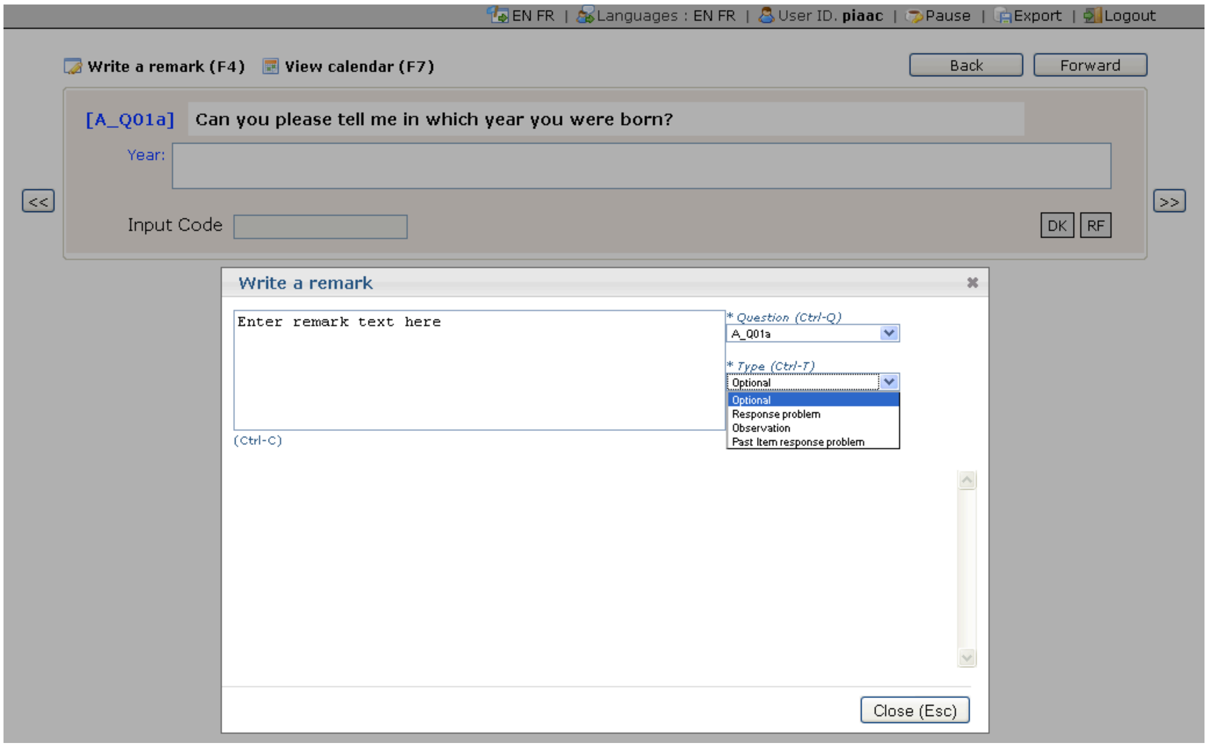
The system checked whether input constraints were met when the interviewer entered inputs with the keyboard. Such verification prevented the interviewer from skipping a mandatory input field or entering values that were not compliant with a mask, a set of acceptable values, a field length limitation, or a numeric range. The check was made on the client side before sending the value to the server. Whenever a constraint was violated when the interviewer validated the response and tried to proceed to the next question, a popup ID displayed explaining the constraint and asking for the answer to be modified accordingly.

After the interviewer introduced the answers of the respondent using the various response fields and codes and pressed the next question key, the system evaluated the consistency checks that might have been specified for that answered question. Whenever the consistency rule was violated, the system produced two types of messages: soft edits and hard edits. Edits are feedback messages presented to the interviewer to report an inconsistent response. They resulted in a pop-up overlaying window with a message to the interviewer identifying the problem and explaining how to fix it. Soft edits identified a range of consistency problems that did not have to be resolved for the interview to continue. In most cases the interviewer was presented with options for resolution that included a means of ignoring or suppressing the edit. In contrast, hard edits identified consistency problems that needed to be resolved for the interview to continue (this included mandatory questions). When edits were resolved, the system checked for the existence of a posteriori inference rules and eventually executed them if necessary.

Once the interviewer instructed the system to reach the next question, if a branching rule was defined, the system proposed the next question according to the branching rule. If not, the system jumped to the next question in the default path, that is, in the path defined by the question definition order.

At any time during the interview, the system allowed interviewers to write down comments about the interface or the question and to modify the comments they wrote. The comment content specified its scope and the diffusion level (private, public, other groups) as shown in Figure 7.7. The system also recorded the event and the timestamp, as well as the author of the comment. In some circumstances, the interviewed needed to be paused by the interviewer and further resumed. Pausing an interview generated a trace in the log for future audit and the interviewer was invited to introduce a comment regarding the interruption.

Figure 7.7: Screenshot of the interviewer comment interface



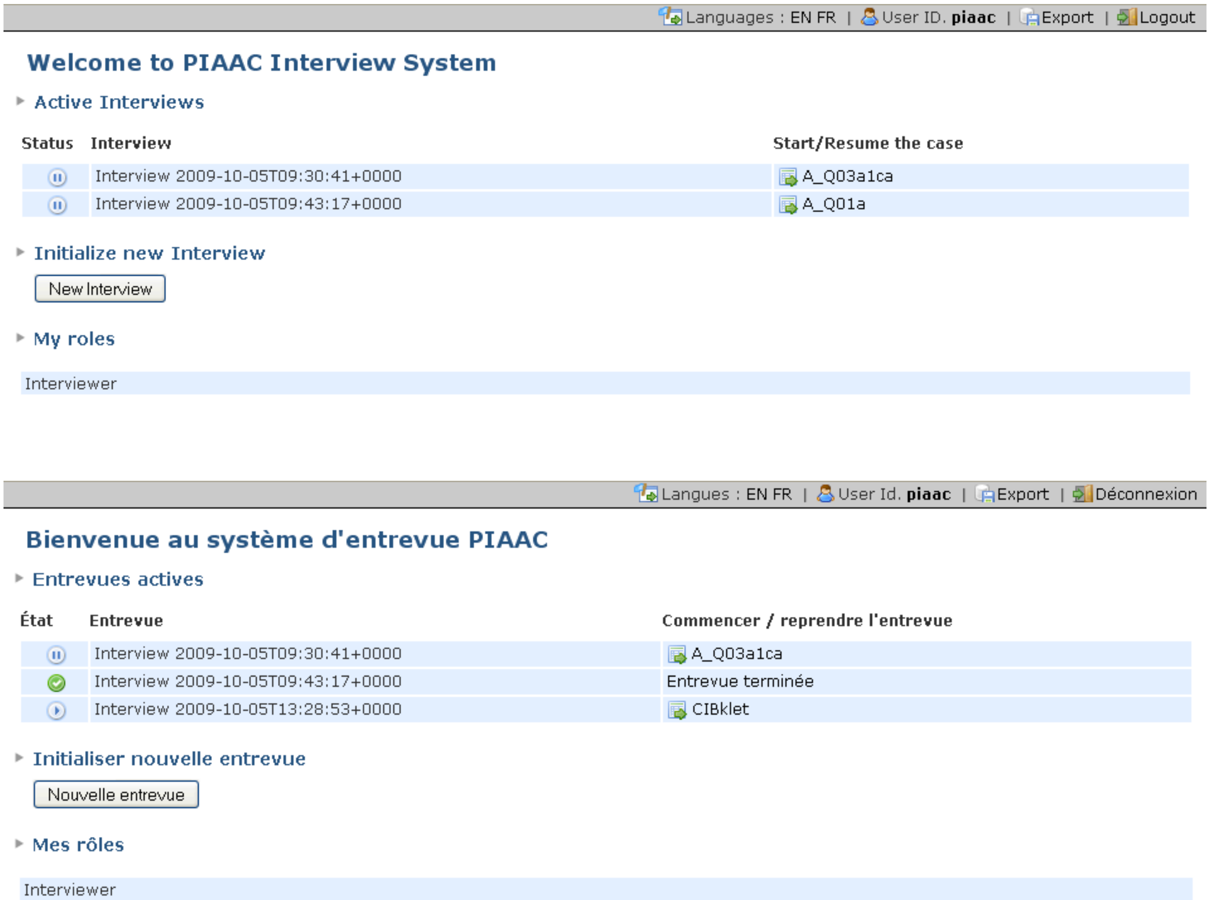
The interviewer also had the option to terminate the interview at any moment. Irrespective of the fact that the questionnaire was terminated prematurely (by the interviewer or caused by a crash) or normally at the end of the question flow, the system assembled all the log information in a series of output files and triggered the export of them to the host system or to the case management system.

7.2.4 Interview administration

Basic administration features were provided in the PIAAC CAPI system, such as a secured password-based interviewer authentication managed at the country level. The interviewers for a given instance of the system were created by the system administrator. The administrator also assigned and scheduled the interviews for particular interviewers.

The administration features of the system also included a basic CAPI system for countries that did not have or wish to use their own. The case management system managed the interview assignment for the interviewers. It showed them pending cases and upcoming ones and let them start, pause and resume their assigned cases. Figure 7.8 shows the interviewer initialization screen, which was made available in different languages depending on the country language profile. The case management system also took care of the interview initialization by importing the initialization file as described previously.

Figure 7.8: Screenshots of the interviewer initialization screen in English (top) and French (bottom)



When cases were completed, it collected the termination codes and comments from the interviewer explaining the context of case termination. These codes were included in the export file of the CAPI system.

In addition to driving the data collection for the BQ, the TAO workflow engine on top of which the CAPI system was built was also used to drive the entire process of the PIAAC survey in households. Indeed, a global workflow was created for the whole process that included the specific one defining the BQ and described in this chapter. Such overarching workflow piloted the survey from the very beginning by sequencing the BQ, the ICT-Core components, the routing of the respondent to either CBA or PBA instruments, and the final section for the interviewer. The ICT-Core components and the CBA instruments were delivered as external activities triggered by the workflow.

7.2.5 Support for PBA instruments

PBA instruments were delivered separately in paper booklets handed over by the interviewer to the respondent. However, some instructions were given by the interviewer at some point, for instance, to instruct the respondent to start a section or turn a page, or to provide him or her with a ruler or calculator. Routing rules were also used to manage the assignment of particular booklets to respondent according to the survey design. To each question of the PBA instrument corresponded a question in the workflow engine where the interviewer had to code whether the respondent provided an answer or not, or refused to answer. Provided that interviewers were trained properly, introducing such code in real time while the respondent was solving the tasks, this protocol allowed for recording timing information for the PBA instrument in the same fashion as for the CBA.

7.3 CAPI runtime

The PIAAC CAPI System Runtime component consisted of an instantiation of the TAO workflow engine running within the VM of the interviewer laptop, which executed a series of embedded workflows and activities made accessible through a Web browser and served by local servers and databases. The PIAAC VM consisted of an operating system fine-tuned for PIAAC operations (lightweight and locked, including only the necessary third-party software and system services), a full TAO platform running under this OS, as well as TAO third-party extensions such as the CBA Item Runtime Environment. The TAO platform contained all the necessary information to run:

1. The global PIAAC interview workflow describing the case initialization, the disposition codes, the ICT-Screener, ICT-Core and ICT-Tutorial, the navigation among instruments, the booklet selection controls, and the interviewer instructions for PBA booklets;

2. The BQ; and

3. The cognitive instruments including general and domain-specific orientations.

Among the cognitive instruments, PSTRE was managed entirely by TAO, while reading and literacy instruments were partly delivered by TAO and the CBA Item Runtime Environment.

While TAO piloted the global survey workflow execution of the study, strictly speaking, the part of the system pertaining to the CAPI essentially covered the way TAO was configured to support the questionnaire definition (ontology and related question definitions) and the questionnaire execution (the interview process and the rule system). In this section, we provide an overview of these core elements of the CAPI runtime.

7.3.1 Components

7.3.1.1 Ontology management and question definition: Generis4

As mentioned earlier, TAO is a CBA platform entirely built on knowledge technologies, and particularly constructed around data structured according to the RDF and RDFS standards (Resource Description Framework, 2004) that represent the foundation layers of the Semantic Web. RDF and RDFS, called RDF/S when considered together, are XML-based languages designed to formally describe ontologies. Ontologies are shared conceptualizations of things that exist in a particular domain of interest (Sowa, 2000; Mahalingam & Huhns, 1997). They describe explicitly the structural part of a domain knowledge in a knowledge-based system using languages with precise primitives (Maedche & Staab, 2001) and associated semantics that is used as a framework for expressing the model (Decker et al., 2000), among which concepts, properties of and relation between concepts, as well as instances of concepts. The formal character ensured that an ontology was machine-processable and exchangeable between software or human agents (Guarino & Giaretta, 1995; Cost et al., 2002). In some pragmatic situations, it simply consists of a formal expression of a metadata framework to describe information units (Kahng & McLeod, 1998).

The kernel of TAO is an ontology management system called Generis4 (standing for GENERIc System for Storing, Structuring, and Sharing knowledge) manipulating RDFS ontologies that structure explicit RDF data and inferred RDF statements according to the RDFS entailment rules, stored in a schema-less relational database. In addition to data-related services, Generis4 also offers a framework to develop Web applications based on Semantic Web technologies, exploiting specific business logics and their associated ontologies in the form of extensions. TAO is such an application based on a series of extensions bearing their own ontologies defining the domain and subdomains of CBA, along with the related application logics and user interfaces. In its basic version, TAO is composed of extensions related to the management and definition of items, tests, test takers, groups of test takers, delivery scheduling, and results export, together with a workflow engine.

As part of TAO, the CAPI system is itself built as an extension exploiting the capabilities of the workflow engine. In the CAPI system, the respondents were considered as test takers bearing a series of characteristics describing the information that the PIAAC survey intended to collect about them and their context – the respondent model and their context form the ontology of the PIAAC survey. Mapping the variables with the characteristic, the collection of the PIAAC data was thus viewed as eliciting knowledge about respondents, that is, a means to instantiate the PIAAC ontology for specific respondents for which precise values were assigned to their characteristics. In TAO, the graphical interfaces to edit the ontologies and their instances were generated automatically from the model. So were the questions displayed in the interviewer CAPI interface, based on a specific question model and templates to display them on screen.

In Generis4, to ease the data and model consultation and editing, users could define specific subparts of the models in the form of hyper classes (they roughly corresponds to views in relational databases). To render the hyper classes for consulting and editing purposes, users could also associate hyper views to them. Hyper classes could be instantiated in the form of hyper instances. A Generis4 service enables one to display hyper classes in edit mode rendered according to its hyper view, and fill in the editable fields to create an hyper instance. In the CAPI system, each question corresponded to an hyper class that was instantiated as a hyper instance for each interview.

In normal use of TAO and Generis4, the different ontologies and their instances were defined using the graphical user interface (GUI) of the system. Performing such knowledge modeling operation requires advanced modeling skills for which PIAAC personnel in countries had low or no competency. In order to facilitate the description of questions and their presentation and execution options, the Consortium provided a specific import/export format of the BQ, together with distinct ad hoc authoring, rendering and management tools. The Consortium took care of importing the files prepared by the countries into TAO.

7.3.1.2 Workflow management and rule system: TAO WF engine

As already explained, the sequence of questions in PIAAC was managed as a workflow where each question represented a data collection activity that fed the corresponding ontology element representing a survey variable. The TAO workflow system consisted of two main parts dedicated to the workflow authoring on one hand, to the workflow execution on the other hand. Depending on country adaptations, the flow of questions might vary significantly between countries. This variability prevented the Consortium from centrally building a single workflow, which resulted in the need for countries to edit their own one. As for the question definitions, defining a workflow was not an easy task and required technical competencies that were not necessarily present in each national team. The Consortium thus provided tools to define the question flow together with the question definitions.

A workflow in TAO was defined as a set of activities sequenced by connectors. Activities are placeholders for services that can be of two distinct types: interactive and noninteractive. The former services correspond to all services provided by TAO through a GUI. Indeed, in addition to the classical GUI, TAO exposes all its atomic management activities, such as item authoring, item metadata editing, and so on as services that can be embedded in workflow activities. Besides TAO functionalities, hyper class services can also be embedded in activities to display forms that feed specific parts of the ontologies. This is the way the questionnaire was executed in the PIAAC CAPI system. Besides TAO services, other interactive services can also be embedded in workflow activities. We used this capability to embed remote item execution from the CBA item runtime in a seamless way for the user.

Noninteractive activities mainly consist of background calculations or calls to distant Web services. In the CAPI system, such activities were used in the process to execute inference rules that precalculate some variables from previously collected data. Noninteractive activities were also used to trigger system functions at relevant moment during the interview, such as the generation of the export file at the end of the interview.

Activities were sequenced by connectors. While in principle, there existed a split (to create parallel flows) and joint (to gather different incoming flows in a single one) connectors, in the CAPI system, none of these were used because the interview only consisted of a single, possibly branched, path. Therefore, only linear connectors bridging one activity to the next and conditional branching connectors were used. The later ones defined possible branches that could be taken according to the evaluation of a logical rule. The conditional branching connectors were intensively used in the PIAAC BQ, as well as to implement the testlet level adaptation of the cognitive instrument. In order to ease the question flow authoring by the countries, the Consortium created a simple rule language that is presented further in this chapter.

7.3.2 Import, export and interface

The communication between any external software, installed on the host system, and the TAO platform, installed on the VM, could be made either by exchanging data or by triggering services.

All data imported or exported by the TAO platform was formatted in XML files. All XML files were always validated with respect to their corresponding schema described in an XSD file. As a consequence, if an invalid XML file was provided to the system, it was not imported and an error was triggered. Imported and exported XML CAPI variable data files were structured according to the same XML schema, described in a XSD file. All data that were either imported or exported are located in folder structure on the host system which was replicated into the VM. An automated mechanism ensured the synchronization of both folder systems. The shared folder structure is described in Figure 7.9.

Figure 7.9: Structure of the shared folder between the PIAAC VM and the laptop host system for file exchange

```
C:\PIAAC\
C:\PIAAC\Platform\
C:\PIAAC\Input\
C:\PIAAC\Output\
C:\PIAAC\Patches\
C:\PIAAC\VMware\
C:\PIAAC\Administration\
```

Only three folders were used to administer the VM, import data and export results during field operations:

1. Output. This folder contained all data collected in the PIAAC survey in XML format suitable for exploitation and analysis. While these data could be further reimported into TAO, they were not exhaustive enough to enable restoring a case or a VM in a given state.

2. Input. This folder contained all data that would be imported into the TAO platform when it was launched. These data were either provided by the National Center and copied into the exchange/import folder, or directly written by a case management system and/or a screener installed on the host system as external software.

3. Administration. This folder contained all data enabling case and full system migration or backup, which preserved the current state of the system. These files consisted of database SQL dumps. They were intended to be used for survey operation management only.

The other folders were used for technical maintenance.

A series of services provided in the PIAAC VM could be triggered from the host system. TAO-related services were triggered using an HTTP request with parameters passed via the URL. These services were made accessible from outside the VM using scripts that could be invoked from the host system either manually, or triggered by an external application such as a case management system. VM administration services were triggered using scripts.

7.3.2.1. CAPI initialization

In many cases, the interview started with a series of predefined information relative to the respondent or the case administration (for instance in the Field Test, some random assignations of BQ parts and assessment instruments were predefined and used to configure the workflows for each case), either generated by the case management system or preloaded by the countries in the interviewer laptops. To enable this, an import function was put in place to parameterize the country-adapted workflows (globally or on a case-by-case basis) and prefill variables from registry data or case management systems. Such imported data were placed in the shared input subfolder, which was systematically read and processed by the TAO workflow engine at startup time.

The input subfolder contained at least one global **initialization** file, and optionally a series of **case-related input archives** identified by the respondent's personal identification (PERSID) number and containing one or more variable files.

The **initialization** file was a unique mandatory file present on each laptop. It contained the country-specific TAO behavior parameters and general workflow control variables that were global from all cases to fine tune the TAO system. It thus represented the default invariant value definitions at the level of a country. It was imported each time TAO was invoked, prior to the case initialization data. This data could, however, be overwritten by case-specific data that might be imported afterward using the case initialization file. The initialization file had the same format as the export variable files to guarantee import and export symmetry, meaning that any exported variable could be further reimported as an initialization one or vice versa.

The **case initialization** was made by importing a series of prefilled variables defined in an archive identified by the PERSID of the respondent. As for the global initialization, the file format of the variable files for import and export was identical. The case-level input archive contained the values of all variables to be imported into TAO. These variables could be any existing variable in the BQ, the general PIAAC workflow, the PBA instructions, or the ICT-Core section. Several of these variables were mandatory to start the interview and were referred as the case initialization in the general workflow. The presence of one or more of these files was checked each time TAO was invoked, after having imported the init.xml file. If present, the PERSID variable XML file corresponding to the current interview was loaded with precedence on those taken from the initialization file. The case initialization files could either be written by the case management/screener external system (depending on countries), or directly uploaded from the NC as part of the case assignment, or collected by the interviewer through other means (mails, CD, memory stick, etc.) and copied into the import folder using a script available on the host system. If no such file was associated to a case, TAO then started with the case initialization section of the general workflow.

The launch of the VM could be triggered either on start-up, manually, or from third-party application running on the host system (case management system, screener, etc.). As a Web application, TAO was launched using an URL that could optionally by complemented by a particular PERSID. The presence or absence of the parameter depicted different situations and induced different file import sequences, as well as different user interactions.

Invoking TAO with no predefined case was the simplest method for countries that did not have a third-party client infrastructure to be executed on the interviewer machine. Such a process occurred in two basic situations. First, when the case assignment was made using lists sent to interviewer on a paper format, such as Word documents or Excel worksheets, no case initialization file preexisted in the input shared folder. Then, an assignment sheet provided to the interview with the minimal required information necessary to initialize a new case using the dedicated section of the interview. In the second situation, the case assignment was made by the NC, which sent (by electronic mail; shipment of physical storage devices such as disks, CD, or memory sticks; by download from the NC-secured website; by network-based folder synchronization, etc.) one or more case initialization files in addition to the global initialization one. The case initialization files then contained the minimal required information necessary to start cases. And depending on the global parameterization, the imported case variable or the missing information could be verified or collected through the case initialization section of the interview as the case was.

Skipping the case initialization by setting the global parameterization variable assumed that the required case variable was imported. Those mandatory variables were the PERSID; the respondent's name, age, gender, address, telephone number; the randomly preselected CBA and PBA booklets (depends on whether the system was configured for the Field Test or the main data collection); and in the Field Test, a variable that controlled the 200 rule (see Chapter 14).

TAO could also be invoked by an interviewer by specifying a specific case. This situation particularly arose in countries using their own case management system and possibly their household screener. The initialization variables were managed by these applications that generated the initialization variable file and launched the CAPI system automatically for the specific case. The Consortium also provided simple case management facilities that enabled countries to preload a series of initialization files for several cases. The interviewer started the TAO CAPI system with an entry point that enabled them to select the case to start (or to resume).

7.3.2.2 Administration data exchange during and after the interview

The administration of the machine relative to interviews consisted of the exchange of database dump files between TAO and the host machine, via the case management system if any, or in any other way countries found convenient. The exchange of the dump file went through the administration subfolder of the shared folder. Two dumps could be triggered at any moment during the interview: full system dump and case dump.

The **full system dump** was made on demand by a component from the host system installed by the countries. Its main purpose was to enable the migration of the full system in its current state from one VM to another VM, from one machine to another machine. The target VM or machine would thus be in the same exact state as the source one. The file contained a full dump of the MySQL database where all TAO CAPI data were stored. The dump also contained all the process definitions (the PIAAC survey definitions), the workflow definition, all cases, their paths and variables. The dump file did not contain the cognitive instrument intermediate results and log files, which were managed outside until the booklet execution are achieved. However, the results of completed booklets were exported within the dump.

The **case dump** was also made on demand from the host system and enabled transferring started cases to a different interviewer in its current state from one VM to another VM, from one machine to another machine. The target VM or machine would thus contain the case in the same exact state as the source one. The file contained a full dump of a particular case from the MySQL database where all TAO CAPI data were stored. As for the full dump, the case dump file contained all the necessary information to retrieve a case exactly in the same state as it had been, except for the cognitive instrument intermediate results and log files of an ongoing case, which were managed outside until the booklet execution ended.

All dumps could also be reimported manually into the VM. This operation was not accessible to interviewers but possible only for technicians at the NC level. Dumping the system for administrative purposes was made using a series of scripts that could be invoked from the host system.

The VM and the instance of TAO installed therein provided a series of functionalities to control the CAPI system. These scripts could be called manually or by other programs on the host system. They provided base functionalities to: start the VM with or without specifying a case, at boot time on the interviewer laptop of triggered by an external application; stop the VM gracefully via normal shutdown of the laptop (crashed VMs could also be stopped to be able to resume normal operation on the next start); start the CAPI system within a running VM with or without specifying a case, stop the CAPI system for a specific case or globally; export results for all or a single case, import all cases that were present in the input shared subfolder or a specific one, dump the cases (all or specific cases) or the whole TAO system; and drop cases for VM cleaning purposes.

7.3.2.3 CAPI result export

Because the TAO workflow drove the entire case, and not only the BQ, the exported files contained all the information pertaining to a case. They were exported in the shared folder between the PIAAC VM and the host system. The export subfolder contained a series of archives, each gathering the output of terminated cases at normal completion of TAO, that is, when the data collection was interrupted temporarily (when the interview was paused) or definitively (when the interview was terminated at the end or before the end) by the interviewer. Each case was exported as a separate compressed file package named according to the respondent PERSID. Depending of the interview completion status, the exported archive might contain zero, three, four or five different files.

The archive contained zero files when TAO exited abnormally and did not export any file. Relaunching TAO triggered the internal recovery mechanism that did not make use of the output files. The partial data could, however, be exported manually using the export services provided in the CAPI system. The archive contained three files: either in case of normal exit from TAO upon premature termination of the interview before or during the cognitive instrument section when delivered electronically; or in case of normal termination of TAO when the respondent was directed to the PBA delivery of the cognitive instrument. The situation when the archive contained four files arose when the cognitive instruments were successfully delivered electronically to the respondent but the post-processing of the raw result file (containing the logs) had failed to export the data properly. Finally, the archive contained five files when TAO terminated normally, with all cognitive instruments delivered electronically to the respondent.

The five exported files per case contained all data collected during the interview and the assessment, together with contextual informations such as the timed log of all events that occurred during the TAO execution and information about the valid and invalidated paths followed during the interview activity flow.

In the archive of a single interview, the unique **Path** file was formatted accruing to an XML schema and contains all information regarding traversed valid and invalid paths along the full PIAAC workflow (including the BQ, and all other steps of the PIAAC survey). The file thus contained the BQ path together with the case initialization, ICT-Core, ICT-Tutorial, ICT-Screener, cognitive instruments and paper-and-pencil instruction related paths. It contained the list of explored question groups (corresponding to the each step of the interview, materialized by a unique screen); for each group of questions, the list of unique questions corresponding to a unique variable; and for each question group, a flag specifying if the explored item group was part of the valid path or not. Interview steps (question groups grouped as atomic activities corresponding to one single screen) covered the whole interview process. If no step of the interview was performed (in case of immediate termination of the interview and normal termination of TAO by the interviewer), the file only contained the topmost XML tags with no question-related information.

The **Variable** file was also unique in the archive of a single case and contained the values of all variables specified in the BQ and the global workflow (according to nationally adapted version specifications) in their final state defined into the TAO platform for a given case. It was formatted according to a specific XML schema and included the data that were imported together, the collected data pertaining to the case initialization, as well as all data collected through the BQ, the ICT-Screener, the ICT-Tutorial, the ICT-Core, and the observations from the PBA instruction sections. The electronic cognitive instrument information was exported in a separated file in the archive. This file constituted the final state of all variables and did not contain intermediate values that might have been changed during the process of the interview. The history of change could be reconstructed from the log file.

In addition to variable names and values, their validity was also provided consistently with the question flow and rules. Hence, all variables that were assigned a value during import or data collection but which finally ended up in a dead branch of the flow, or that were never addressed during the interview, were flagged as invalid. At export time, TAO did not clean the data, enabling the widest range of post-processing possible. In case there were variables for which no value was imported or collected, they were reported using an XML tag with variable name but no value attribute. When the variable file was used for import, if no value attribute was provided, the variable was ignored. On import, all validity attributes were ignored.

The **log** file was formatted according to a specific XML schema and contained the trace of events occurring both at the server side (response of the system to user requests) and the client side (actions of the user on the interface) pertaining to the case initialization, the BQ, the ICT-Screener, the ICT-Tutorial, and the ICT-Core related items, as well as the PBA instructions. All entries in the log were timestamped and enabled to reconstruct the entire sequence of user action and system responses of the CAPI system. An equivalent logging of events was provided within the cognitive instrument result file for further analysis at psychometric level. Except for the cognitive instruments (which were operated by the respondent), all actions of the interviewer were recorded in the log. This file was also used to generate audit trails for verification of interview quality.

Cognitive instrument results were exported in a series of separate files corresponding to clusters of CBA instruments for a given case. The result consisted of the scores (when scoring was made automatically) for all tasks, units, and tests, the PERSID of the respondent, the responses, and other contextual information. It also contained the log of all recorded user actions on the interface. The result file was exported at the end of each booklet execution, as part of the assessment workflow service, before the PIAAC general workflow automatically resumed and proceeded to the following survey steps. Crash recovery data were not stored in this file but in a dedicated structure enabling recovery independently of result export.

The raw cognitive instruments result file was seconded by another post-processed file containing the **CBA Variables** only. The supplementary result file contained the same information as the previous one, with the exception of logs, which were post-processed to provide meaningful additional scoring-related variables, instead of lists of atomic events. It was exported at the end of each booklet execution by the assessment workflow service, before the PIAAC general workflow automatically resumed and proceeded to the following survey steps. It was produced by a post-processing routine that analyzed patterns of events in the raw result CBA file and generated a series of variables that were used for further analysis and scoring (in the case of PSTRE).

The generation of export files was triggered at each normal completion of TAO, irrespective of the case status (completed or not completed). When an interview was started, it could be paused or terminated with a complete case or with a partial case (definitive termination before reaching the end of the interview). When the interview was paused, the result files were exported with the disposition codes relative to the three main survey sections: BQ, ICT-Core and cognitive instruments. In the presence of a case management system, the exported codes could be used in monitoring processes set up by the countries. Whenever the case was prematurely terminated by an interviewer operation, the workflow proceeded to the last activity of the flow relative to the section, ensuring the interviewer was presented the correct disposition code entry screen before ending the interview. At the end of the interview, whether upon premature termination or normal termination with a complete case, the result files were also exported. Resuming a case with the sole exported results was not sufficient. Such operation required other information that was stored in the database.

References

Cost, R. S., Finin, T., Joshi, A., Peng, Y., Nicholas, C. Soboroff, I., . . . Tolia, S. (2002). ITalks: A case study in the semantic web and DAML+OIL. *IEEE Intelligent Systems*, *17*(1), 40-47.

Decker, S., Melnik, S., Van Harmelen, F., Fensel, D., Klein, M., Broekstra, J., . . . Horrocks, I. (2000). The semantic web: The roles of XML and RDF. *IEEE Internet Computing*, *15*(5), 2-13.

Guarino, N., & Giaretta, P. (1995). Ontologies and knowledge bases: towards a terminological clarification. In N. Mars (Ed.), *Towards very large knowledge bases: Knowledge building and knowledge sharing* (pp. 25-32). Amsterdam, Netherlands: IOS Press.

Kahng, J., & McLeod, D. (1998). Dynamic classificational ontologies: mediation of information sharing in cooperative federated database systems. In M. P. Papazoglou & G. Sohlageter (Eds.) *Cooperative systems: Trends and direction* (pp. 179-203). San Diego, CA: Academic Press.

Maedche, A., & Staab, S. (2001). Ontology learning for the semantic web. *IEEE Intelligent Systems*, *16*(2), 72-79.

Mahalingam, K., & Huhns, M. N. (1997, June). An ontology tool for query formulation in an agent-based context. *Proceedings of the Second IFCIS International Conference on Cooperative Information Systems.* Kiawah Island, SC, 170-178.

Resource Description Framework (RDF) (2004). Retrieved from http://www.w3.org/RDF/

Sowa, J. F. (2000). *Knowledge representation: Logical, philosophical, and computational foundations.* Pacific Grove, CA: Brooks Cole Publishing Co.

Chapter 8: Development of the Integrated Computer Platform

Britta Upsing, Frank Goldhammer, Maya Schnitzler, Robert Baumann, Roland Johannes, Ingo Barkow and Heiko Rölke, DIPF; Isabelle Jars, Thibaud Latour, Patrick Plichart, Raynald Jadoul and Christopher Henry, CRP; and Mike Wagner, ETS

This chapter describes the electronic test delivery system of the PIAAC study. The complete delivery system consists of three parts that were installed on the interviewer's laptop. First, the country's case management system selected and organized the participants in the study and interacted with the embedded virtual machine (VM). As this was country-specific software, we cannot specify the content of these programs, but we can state that a common interface was defined, making access of the PIAAC VM possible. The second part was the VM itself, which served as an encapsulated environment to installed software from other household surveys and to the different hardware configurations existing in different countries. Third, there was the BQ and cognitive items running on the TAO platform within the VM. All these parts interacted; the workflow of interaction is described here in Figure 8.1.

Figure 8.1: Design of the electronic test delivery system

1) Interviewer Laptop: *Host System*
- Country-specific software
- Country's case management system
- VMware runtime

2) DPVM: *Guest System*

3) *TAO Platform*
- BQ
- Cogn. Items

The interviewer laptop ran its normal operating system, the *host system*, which was supposed to be a Windows system (see the technical standards documents). The host system ran optional country-specific software such as a country's case management system and the runtime for the virtual machine.

The virtual machine (DPVM: delivery platform virtual machine) ran within the host system. It is called a *guest system*. It ran TAO, including the BQ and the cognitive items.

The delivery of the BQ and cognitive instruments was done by means of TAO, which ran within the VM. In doing so, there were a minimum of dependencies, influences and interferences between the PIAAC delivery system and the interviewer laptop.

8.1 Development of the virtual machines

The VMs containing the PIAAC system and items are described here in more detail.

8.1.1 VM basics and hardware

All VMs delivered in PIAAC were based on the same prototype VM system. The technical details are as follows:

- VMWare Workstation 6.5-7.x virtual machine
- Virtual hardware
- 1 GB RAM
- Single core processor
- 40 GB HDD
    - Dynamic allocation
    - 2-3 GB actually used
- Display 1024 x 768 pixel

8.1.2 Content of the PIAAC VM

The operating system running inside the virtual machine was a Debian Linux system. Debian ensures a high level of dependability combined with the assurance that no licensing is necessary (open source strategy). Unnecessary software of the standard Debian distribution was removed to save space. This is an overview of the software components used in the PIAAC VMs.

Operating system:

- Debian Lenny
- Kernel 2.6.26-19 – Aug 2009
- X Windows
- IceWM Window Manager

LAMP stack

- Apache2 2.2.9-10
- PHP5.2.11-0.dotdeb.1

- Suhosin-Patch 0.9.7
- Zend Engine v2.2.0
- MySQL v14.12

8.1.3 Development process and automation

There are different parts of a VM. The base is the “Mother VM.” It contains all common data, like the operating system and all software components, which were independent of what country used it. This VM needed to be supplied with all country-dependent information and data.

The Item Management Portal (IMP) held this information. Each country had an authorized person adopt items and translations.

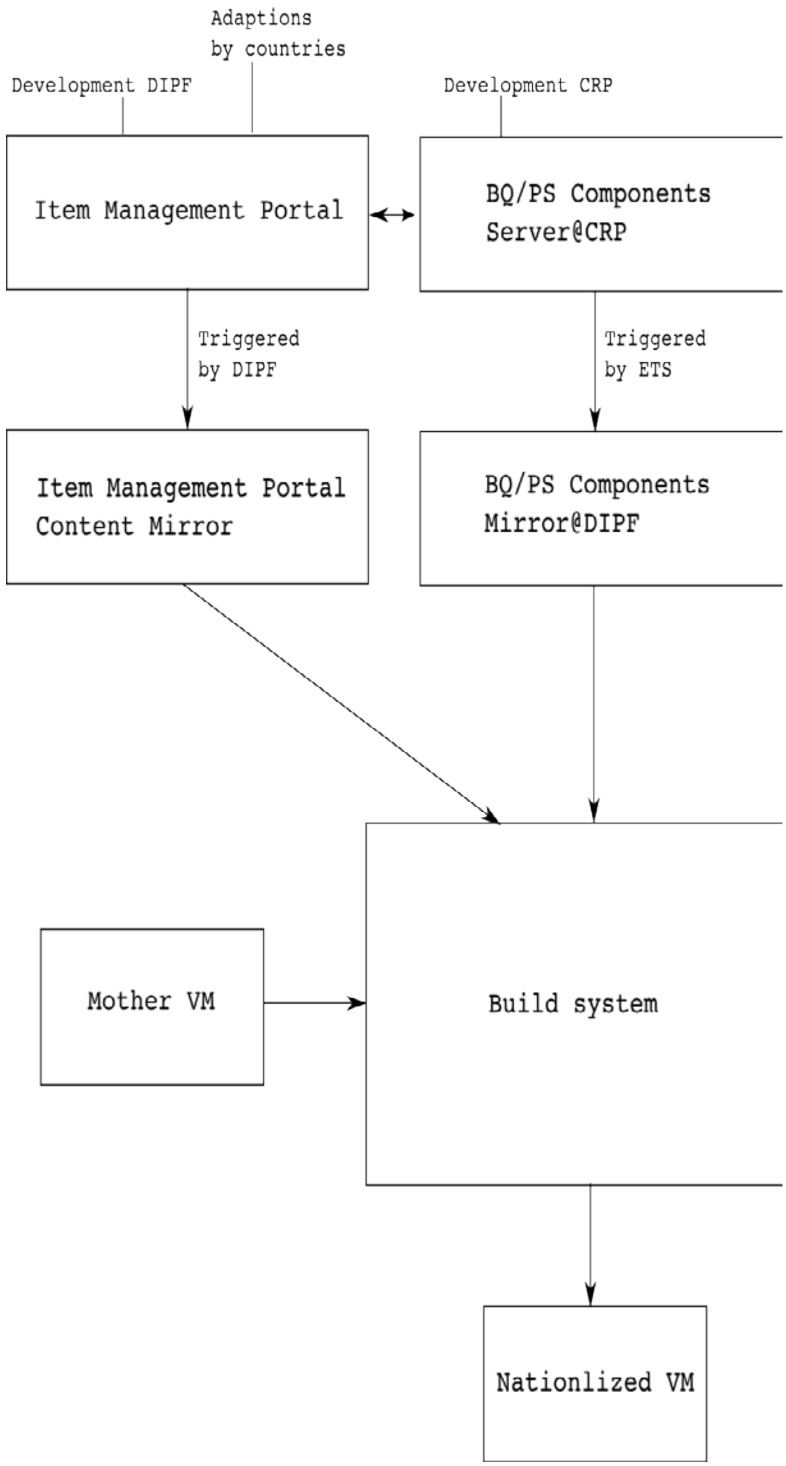
Several servers from the Consortium partners provided all necessary data to the IMP.

At dedicated points of time or after important updates, two main servers were mirrored for the build process. This helped get dedicated versions with timestamps.

The build system fetched all data from the mirror sites and combined this with the Mother VM.

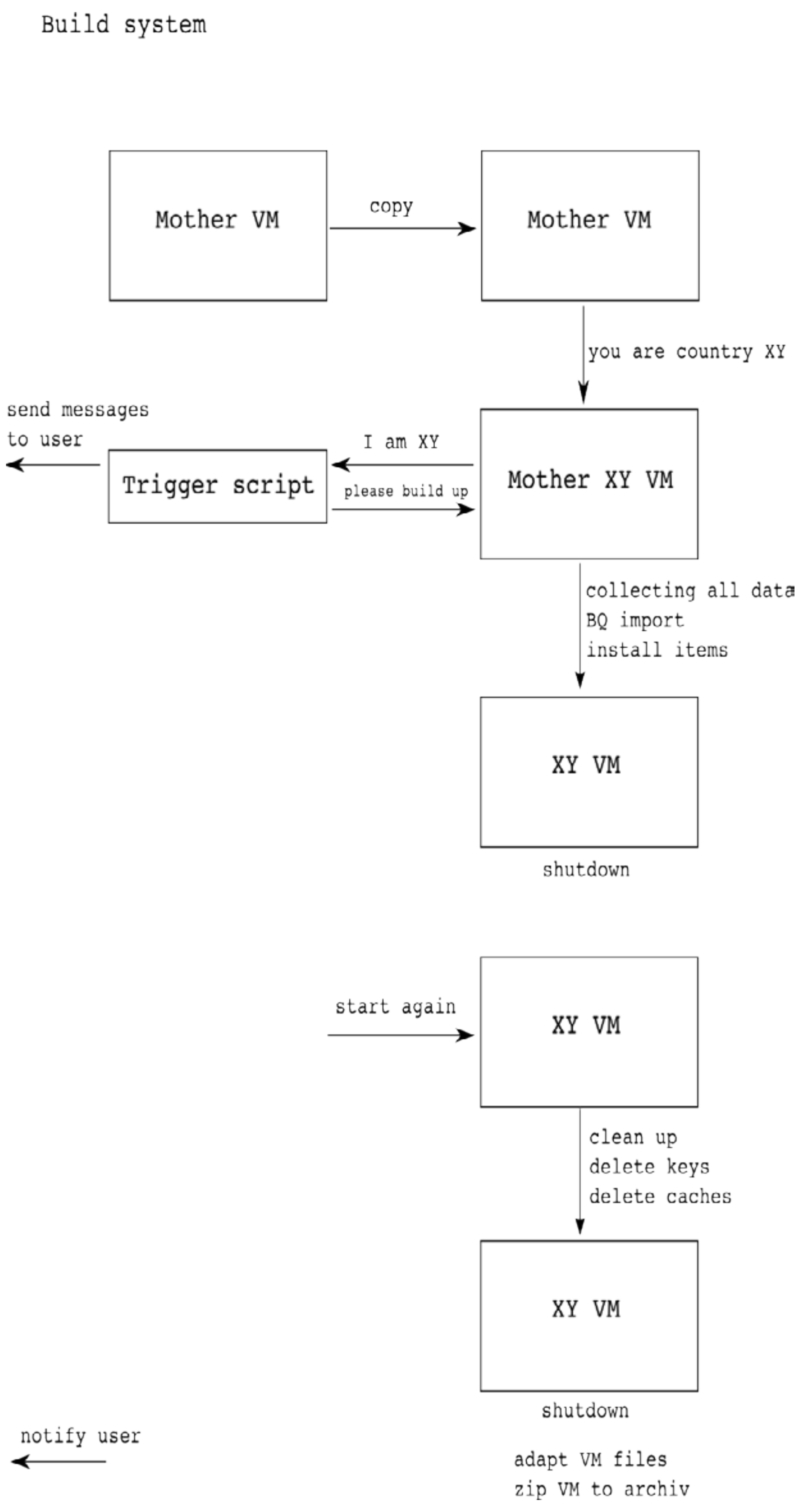
This process is denoted in Figure 8.2.

Figure 8.2: Development process and automation



The build system (Figure 8.3) was a collection of scripts, based on the Bourne again shell (bash), on a build host.

Figure 8.3: Build system



The process was triggered by a bash script for mass production. A loop would call for one task, which will be described as follows. A template of a Mother VM was copied to a new folder. The build system started this copy. The VMware VMPlayer software needed an environment to place a window. A virtual frame buffer was installed and all graphic output was dropped there. It is important to know there was no user interaction possible or necessary. Installing a new Mother VM needed a start by hand to ensure there were no error windows blocking the process. The VM received a country ID as its name in the configuration file. The VM reading this environment variable and contact a script outside via secure shell (ssh).

This trigger script controlled the process from outside. The advantage of this concept was that the messages of the inner scripts were passing the trigger script and could be sent to the originating user. The trigger script sent a start mail to the originating user and called a build script inside the VM. The IP of the VM was transmitted when the trigger script was invoked. This inside script was responsible for several tasks. It fetched all data from the mirrors, did a BQ import and configured all locale settings (keyboard, fonts, etc.)

After this, the VM performed a shutdown, and build task started it up again. This second start invoked the cleaning scripts inside. This was done after a reboot, providing the chance to suppress the second boot up for debugging purposes.

Before the VM did a second shutdown, all installation scripts were disabled and a reboot brought up a nationalized VM.

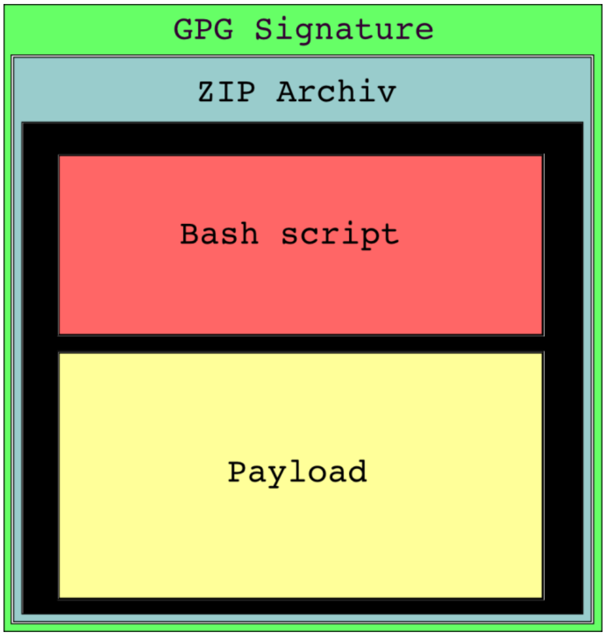
After shutdown, the build script adapted the VM files. Inside the VMX configuration file, some parameters were changed and the caches and the snapshot configuration file were deleted. Four files remained in the VM subfolder. The last step was to zip the folder and send a notification to the originating user.

8.1.4 VM patching process

The intention behind the patch mechanism was to have a robust way to modify the system in any manner (see Figure 8.4). Maybe it would even be necessary to modify the patch mechanism itself. To install a patch, you needed to place a file at a dedicated place inside the VM. This could be done with a graphical user interface provided by the Windows-scripts managing software or via a command-line interface with secure copy protocol (Linux, use WinSCP or the country-specific Case Management System (CMS) Tools under Windows).

The “init” scripts, which were called in the boot process of Debian, had to look for patch files. For these purposes, the init script “/etc/rc.local” was modified. If it found files in “/var/www/piaac/Exchange/patch,” it then checked each file with a GPG signature, unpacked all files inside each zip file and executed a bash script, found inside each patch package.

Figure 8.4: VM patching process



A patch file contained an executable bash script and a payload. The payload consists of additional files that represent the fix itself. These were copied to the destination folder by the bash script. This mechanism allowed many things to occur inside the VM such as kernel changes, exchanges of tests, updating software and many more. This involved considerable risk if you provided a patch file with dangerous contents. There was the possibility of completely destroying the VM. To ensure applying a correct patch file, it carried a correct GPG signature. The keys for signing were different between the Field Test and Main Study, so you could not install a Field Test patch in a Main Study VM. Also it was impossible to install a patch from anyone but the PIAAC Consortium.

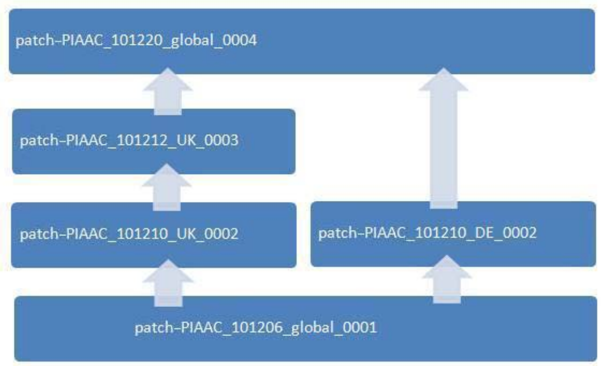
After executing the patch bash script, the patch file was moved to the installed subfolder. The next reboot would not be able to see old files and try to execute them.

For building new patches, a patch-build VM is useful. This could be an "old" Mother VM with the complete subfolder structure and all keys for signing. It was possible to build patches under Windows, but the end-of-line characters in text files were different between Linux and Windows. The patch mechanism would be confused by Windows text files. If you used "Notepad+" for editing, you could configure this feature to overcome the problems. Other issues were ownership and rights of the files inside a patch. Windows users needed to ensure the files would be owned by root and that the bash script had execution rights. It was recommended to use Linux or a patch-build VM for highest efficiency.

The content of a patch file was simple. It contained at least two files. The bash script needed to be called "execute.bash" and have a file called "description.txt". The bash script is explained above; the description.txt file contained a short description of the function – the involved files and fields about target countries, provider, builder and date. The content of description.txt was shown after reboot of a fresh patched VM in a window. This was managed in the IceWM startup file.

The patches were created in a numbered order. Because one patch could be dependent on another, it was necessary to install patches in a correct order.

Figure 8.5: Naming of patches



8.1.5 Changes from Field Test to Main Study

The build chain for the VM was only bug-fixed and some minor changes were done. More modifications were applied to the patch mechanism. To avoid the potential problem caused by trying to ensure patches were installed in the order they were released (as during the Field Test phase), for the Main Study there was no need to take the order into consideration.

For the Main Study, every new patch included the previously released ones (cumulative patch, see Figure 8.6). The VM would know about the so-called patch level, which meant that every VM held a list that contained the patches already installed. Patches already installed were skipped during the installation process.

National patches were included in the cumulative patch. A minor "problem" with national patches was that they only increased the patch level and did nothing else for countries that were not involved.

The naming of the patches for the Field Test was handled individually. For the Main Study the naming was as follows (see Figure 8.5):

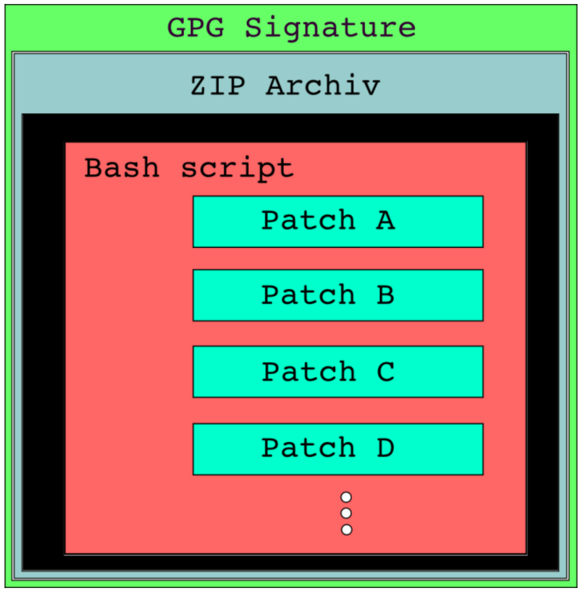
patch-PIAAC_<date>_<responsibility>_<number>.zip.gpg

where <**date**> : the release date, format : yymmdd

<**responsibility**> : was *"global"* for patches to be installed by all countries or will be the country code for national patches (e.g., "*UK*")

<**number**> : the release number starting at "1", format nnnn. The number was increased incrementally each time a patch was released.

Figure 8.6: New design of patch mechanism in Main Study



An exception was that national patches released at the same time for the same purpose but for different countries had equal numbers. For example:

- Both national patches released on 10 December 2010 included the global patch released on 6 December 2010.
- The national patch released on 12 December 2010 included the national patch for England/Northern Ireland (UK) released 12 December 2010 and the global patch released 6 December 2010.
- The global patch released on 20 December 2010 included all national and global patches below.

It was possible that during a phase of PIAAC the cumulative patch file would get too big because there were too many files to replace (e.g., computer-based assessment projects for all countries). It wasn't expected to happen and did not happen, but still there was a strategy to handle this case:

If a patch needed to replace a large number of files or huge files, these files were provided in an external signed file. The user needed to download the patch file and the supplementary file. So this file was required only once. Subsequent patches did not include the supplementary file's contents but would require it if the previous patch wasn't already installed.

The patch itself was extended into the Main Study. In simple words, it contained one patch as with the Field Test, which contained all other patches. The superior patch executed a bash script as usual. But this script coordinated the installation of the contained patches. It checked the patch level and decided which patches should be installed. For each contained patch, there would be the normal patch installation mechanism operating. Each contained patch had the same structure as a Field Test patch (execute.bash, description.txt and payload). A cumulative patch was a meta-patch, which executed all other patches inside in a correct manner. In other words, a Main Study patch was a Field Test patch, and the payload was a bunch of patches.

8.2 Development of the interfacing software for the integration in a national CMS (DIPF/ETS)

The VMs running the PIAAC software were well insulated from the surrounding host system. In fact, this was the main requirement leading to choosing VMs as the building block of the PIAAC delivery. Nevertheless, restricted communication between host and guest (VM) had to be possible, for example, to (re-)start or stop the VM and exchange data such as results or patches. Normally, a so-called case management system (CMS) ran on the host system supporting the PIAAC interviewer by managing the sample and the interview status. The CMS could use the interfacing software as a kind of remote control for the assessment software in the VM.

8.2.1 Interface software requirements

Requirements resulting from the environment were as follows:

- Because Windows (XP, Vista, 7) was used, the interfacing software was to be designed to run under these operating systems.

- The participating countries did not use a common CMS, if any. This required that the software be accessible from any CMS software and not require a special runtime environment.

- The PIAAC VM could be run by three different VMware products, so the interfacing software needed to support these products :

    - VMware Workstation (mainly used for testing and setup purposes)
    - VMware Server (used in the Field Test)
    - VMware Player (used in the Main Study)

The functional requirements for the interfacing software were:

- VM remote control
    - Starting the VM
    - Terminating the VM
- Assessment control and data access
    - Start a case
    - Resume a case
    - Getting a case state
    - Retrieving the result data
- Handle maintenance and administrational requests

- Install patches
- Archiving and recovery

8.2.2 Implementation of the interfacing software

The interfacing software was developed using “AutoIt,” which is a freely available programming environment for Windows. All releases of the software also include the complete source code. This enabled the countries to make changes to the source code if necessary without the need to buy expensive programming tools.

The functionality mentioned above was developed as small programs – the so-called PIAAC scripts. These scripts could be run from the command line or could be called by a program (e.g., a CMS).

The scripts interacted with the PIAAC VM via the VMware VIX interface software. The VIX software provided services such as controlling VMs, file handling and calling programs and scripts inside the VM.

8.2.3 Setup

To install the PIAAC scripts, the scripts were downloaded and copied to a particular location on an interviewer computer.

To handle different user environments concerning the location of the files (scripts and virtual machine) and settings like VMware access information, there was a configuration file which could be modified by a configuration script.

The required folder structure was:

- C:\piaac
  holds the PIAAC scripts (this folder can be changed)
- C:\piaac\input
  holds the prepared input files containing case data
- C:\piaac\output
  holds the result files of finished or paused cases
- C:\piaac\administration
  holds supplementary files like database dumps
- C:\piaac\patches
  holds patches to be applied to the virtual machine

8.2.4 VM remote control

To control the VM remotely, two scripts were developed. Script StartVM, which was developed first for the Field Test, started the VM. This functionality moved to the script HandleCAPI, which was developed later for the Field Test. This script started the VM only if it wasn’t running yet and couldn’t start an interview.

The script StopVM forced the Debian operating system to shut down and terminate the VMware software.

8.2.5 Assessment control

For the Field Test script, StartCAPI handled the start of a case and script ResumeCAPI handled resuming a paused case. Both functionalities in the Field Test later moved to script HandleCAPI, which was able to determine the state of a certain case and was dependent on the state to start or to resume a given case.

Starting a case required an input file on the input folder. This file was copied inside the VM, and after that, a service script inside the VM was called that started the browser and displayed the interview at the state specified by the input file (new interview) or at the state it was paused (already started).

A CMS then could look for the state of the running interview by frequently calling the script GetCaseState. Every call of GetCaseState produced a file containing some data about the state and the progress of the current interview. This file was located on the administration folder. If the state changed from “running” to “paused” or “finished,” a CMS could react accordingly:

In case an interview finished, the file containing the collected interview data could be copied to the output folder by calling script ExportResult.

A CMS then could start or resume a new interview or terminate the VM by calling script StopVM.

8.2.6 Data access

The script ExportResult was used to retrieve the collected data of a certain interview or all interviews. The file or the files was copied from the VM to the output folder. Inside the VM, the file was moved to an archive folder. This was to avoid having a result file copied more than once.

Script ControlCAPI frequently called the script ExportResult, so every time a result file was written, it was copied automatically to the output folder.

8.2.7 Maintenance and administration

Dropping cases
Deleted the interview specified by the PERSID from the TAO database. This operation was irreversible and needed to be used with care. There was no recovery for this action.

Recover case data
Recovered the result files of a certain case or of all cases from the database or from archive folders and copied the files to the Windows environment. The files were not copied by default (see below). Recovering from the archive folders was preferred to get complete result files.

Dump
To get a dump of the SQL database containing the case data of a certain case or of all cases, there were two scripts named DumpCase and DumpAllCases. The SQL dump files were copied to the administration folder.

To import previously generated SQL dump files to the SQL database inside the VM, there were two scripts named ImportCase and RestoreAllCases. The DisableKeyF9 disabled the debug feature “Watch window.” After a VM was tested successfully, this script needed to be executed one time on a clean VM to turn off the debug feature. Note that the feature could not be enabled again.

Patches

If a bug in a VM was discovered up to now, the bug was fixed and a new VM was provided for download. Downloading a new VM meant it would take time until the new VM was available on the target system. Also, because VMs are so large, it could be problematic to deploy new VMs once interviewers had started work in the field. An easier and faster way to fix a bug in a VM which was already installed was to provide just the changes needed in the form of a patch.

A patch is a small file that is provided for download and contains only the changes to fix certain bugs. Because only the changed files, not the entire VM, were included in the patch, transmittal and installation was fast. This saved time. Installing a patch was secure because the files were signed. A VM would not accept a patch file that wasn’t signed or was signed with an invalid key.

There were two methods for installing patches, either via the command line or via a graphical user interface (GUI). For manual patch installation, the GUI version was easiest. For automated installation, for example, via a CMS, the command line version was best.

8.2.8 Interface overview

Scripts were to be called from the command line in a Windows environment. Some of the scripts required one or more parameters, for example, the PERSID, to specify a certain case. One thing to avoid was starting a script by double clicking on the name, for example, in the Windows Explorer. In that case it was not possible to enter parameter values.

Every script displayed a small piece of information about its purpose and usage if the parameter value “**help”** was entered, such as “**StartCAPI help.”** To manipulate the behavior of a script, switches could be added at the end of the command line (last parameter value). In case of more than one switch, the first character was ‘-’ followed by the switches’ names (usually one character), such as “**-dop**”.

All scripts supported the switch “**-d**”, which produced an editor window for debug output. This debug output could be saved to a file for bug report purposes, such as via the bug tracking mechanisms established.

Configuration tool

- PIAACscriptConfig

Basic scripts

- StartCAPI (optional: “new” “login” or PERSID)
- ResumeCAPI (PERSID)
- ExportResult (optional: PERSID)

- ControlCAPI
- StopVM
- GetCaseState (optional: PERSID)
- HandleCAPI (optional: “new” “login” or PERSID)

Administrative scripts

- DropCase (PERSID)
- DumpCase (PERSID)
- ImportCase (PERSID)
- DumpAllCases
- RestoreAllCases
- RecoverCase (optional: PERSID)
- DisableKeyF9
- PatchVM
- PatchVM_GUI

8.2.9 Changes and enhancements from Field Test to Main Study

During the Field Test it was important to apply the patches in the right order. For the Main Study, the method was changed to make the patch process safer.

Patches were now cumulative, meaning each new release of a patch also contained the patches released before. Each patch now owned a so-called patch level. The patch level was a number that was increased incrementally for every new release of a patch. The patch level was part of the file name (see the last four characters).

If a patch was applied to a VM, it compared the patch levels of that patch and the one within the VM. The patch would be applied only if the patch level was higher than that of the VM. Otherwise it was ignored. It was recommended to apply the patch having the highest patch level if there was more than one patch applicable.

The script PatchVM_GUI released for the Main Study offered only patches having a higher patch level than the current VM.

Chapter 9: The TAO Platform

Raynald Jadoul, Patrick Plichart, Jérôme Bogaerts, Christophe Henry and Thibaud Latour, CRP Henri Tudor

TAO is a platform developed as open-source software. It was initially designed for national education monitoring in Luxembourg, Germany and Hungary. It is also used for many other pilot studies worldwide. The major strengths of TAO reside in its flexibility and a design oriented toward a multilingual, highly distributed and cooperative operationalization of survey processes led in an international context.

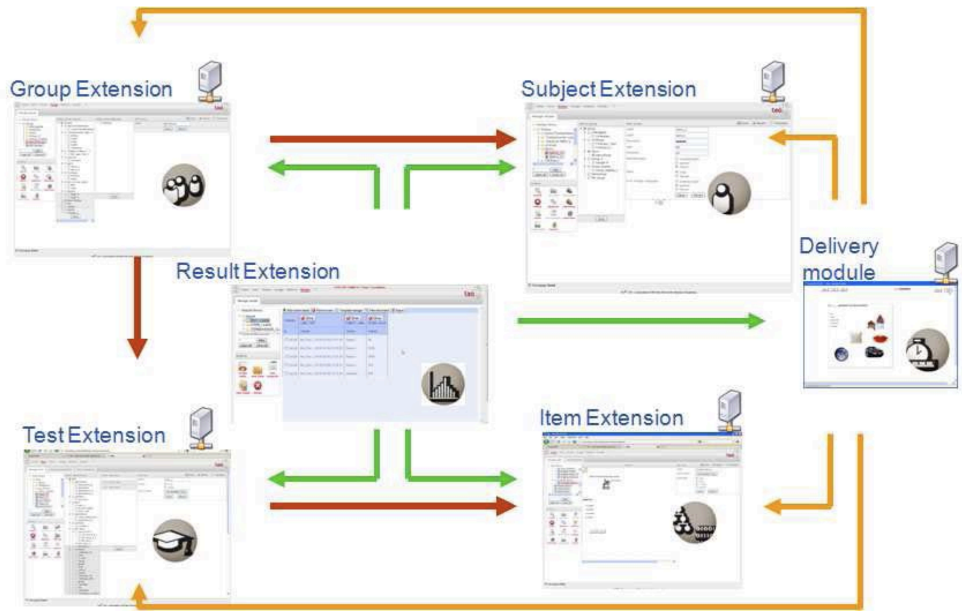
Since 2006, OECD has relied on TAO for a progressive computerization of its large-scale studies (e.g., the Programme for International Student Assessment, or PISA, 2009; PISA 2012, PISA 2015, as well as PIAAC).

As a technical solution, TAO provides a general and open Web-based architecture for the design, development and delivery of computer-assisted tests. It is released under the GPLv2 license and available on the TAO website (http://www.taotesting.com). Although TAO provides much of the functionality required for the operationalization of large-scale assessment processes (authoring tools, workflows for the management of the activities related to the creation of test takers, deliveries, etc.), it was essential to enrich the platform for PIAAC with various features described below.

9.1 TAO architecture overview

In terms of architecture, TAO is built on top of a knowledge base (i.e., a database capable of handling highly flexible data models) called Generis. The TAO platform is 100% Web-based. Its architecture entails independent components (called "extensions") covering the operations involved in a typical CBA lifecycle (see Figure 9.1).

Figure 9.1: TAO's extensions-based architecture for CBA lifecycle



9.1.1 Item (or question) management

This part manages the creation and design of items. An item can be an exercise, that is, a competency assessment that takes the form of a question for which there is at least one right answer and a scoring algorithm related to the competency to be evaluated. Exercises are also used for formative purposes. An item can be an informational question, meaning a question that does not seek to evaluate a competency but to elicit background information that provides the context (social, economic, etc.) of the test taker. Different design templates are proposed depending on the type of item to create.

9.1.2 Test (or questionnaire) management

This part manages the creation of tests. Tests combine a selection of items into a defined set. Test parameters include item order, scoring, layout, and so on. In the terminology of TAO, a test that integrates solely competency assessment exercises is called a competency assessment (or a computer-based assessment; CBA), and a test that is composed uniquely of informational questions is called a background questionnaire (BQ).

9.1.3 Test taker (or interviewee, or respondent) management

In this part, one can register test takers for the platform, define their registration data (e.g., login, password, mother tongue, location, etc.), and associate them with the relevant group(s) to which they belong.

9.1.4 Group of test takers (or interviewees) management

This part manages the creation of groups for organizational purposes. For example, grouping test takers according to global features and classifications (like the citizenship to a country) is managed in this component.

9.1.5 Delivery (test taking, examination, or interview) management

This part manages test deliveries. The creation of deliveries is the process of assigning selected tests and selected groups of test takers to delivery campaigns. During the creation of a delivery, it is possible to exclude test taker(s) on an individual basis, notwithstanding their group membership. A delivery campaign features many parameters: sequence of the tests, maximum number of executions, delivery period, and so on.

9.1.6 Results management

This part manages the results of the "executions of the deliveries" (also called instances); these data include the information of all delivered tests, their related test taker-, group- and item-specific data, as well as the individual data collected during the "test execution" (also called runtime). Individual data collected during runtime include behavioral information (e.g., reaction times, latencies, hesitation) and contextual information (e.g., hardware settings). In this extension, one can create tables to visualize the results and export them for further data analysis.

9.1.7 Process management

The focus of this extension is the creation and operation of the processes (operated through workflows) required to drive the various types of activities needed in the development of large-scale surveys, hence to support all specialists involved in carrying out those activities.

9.2 TAO architecture for PIAAC

The context of PIAAC required building new extensions and support tools on top of the TAO platform.

First, TAO integrated a new type of item to support the CAPI survey style. CAPI surveys do not deliver the question items directly to the respondent; instead, professional interviewers read the questions to interviewees; questions are backed by instructions and complementary information that guide the interviewers. From a technical point of view, this new form of question item was a major challenge; it had to transpose all the requirements and expected capabilities of standalone (i.e., not Web-based), highly responsive, fully-fledged commercial CAPI solutions to a free-of-charge, full-edged Web-based CAPI platform; for example, professional interviewers make intensive use of keyboard shortcuts and need adapted user interfaces. Also, the transition delay when leaving one question for the next should not exceed one second.

TAO was also enhanced to support highly complex flows of items where questionnaires make an extensive use of a) dynamic sequencing of questions (i.e., different answers lead to different follow-up questions), b) dynamic layout of questions (e.g., German words tend to be longer; white spaces are seldom in Japanese; Asian characters need to be magnified by 10% to be readable; Hebrew and Arabic are read from right to left), and c) dynamic wording of questions (e.g., some language grammars require the wording of a question to take account of the respondent's gender, age, and so on). The TAO toolbox integrated tools and processes allowing those dynamic aspects of the questionnaires to be tailored at the national level for each country participating in PIAAC.

Finally the number of partners involved in operationalizing the questionnaires and the diversity of the tasks to be achieved, often sequentially, sometimes concurrently, meant the need for well-defined organizational processes. These processes entailed the operations to create items, adapt them to reflect national specifics, check their quality, push them to the final support (in PIAAC's case, a virtual machine), and so on. Therefore, the "process management extension," based on an ontology-driven "workflow engine" (the first of its kind), had to be enhanced to handle a large quantity of active parallel processes.

In terms of software architecture, the workflow engine was not only responsible to support operationalizing questionnaires. The same workflow engine led the sequence of the noncognitive question items during the interviews. Initially defined by the specification document describing the PIAAC general workflow of the BQ, the sequence was adapted on a national basis.

The TAO architecture entails these **key features**, described in the sections below:

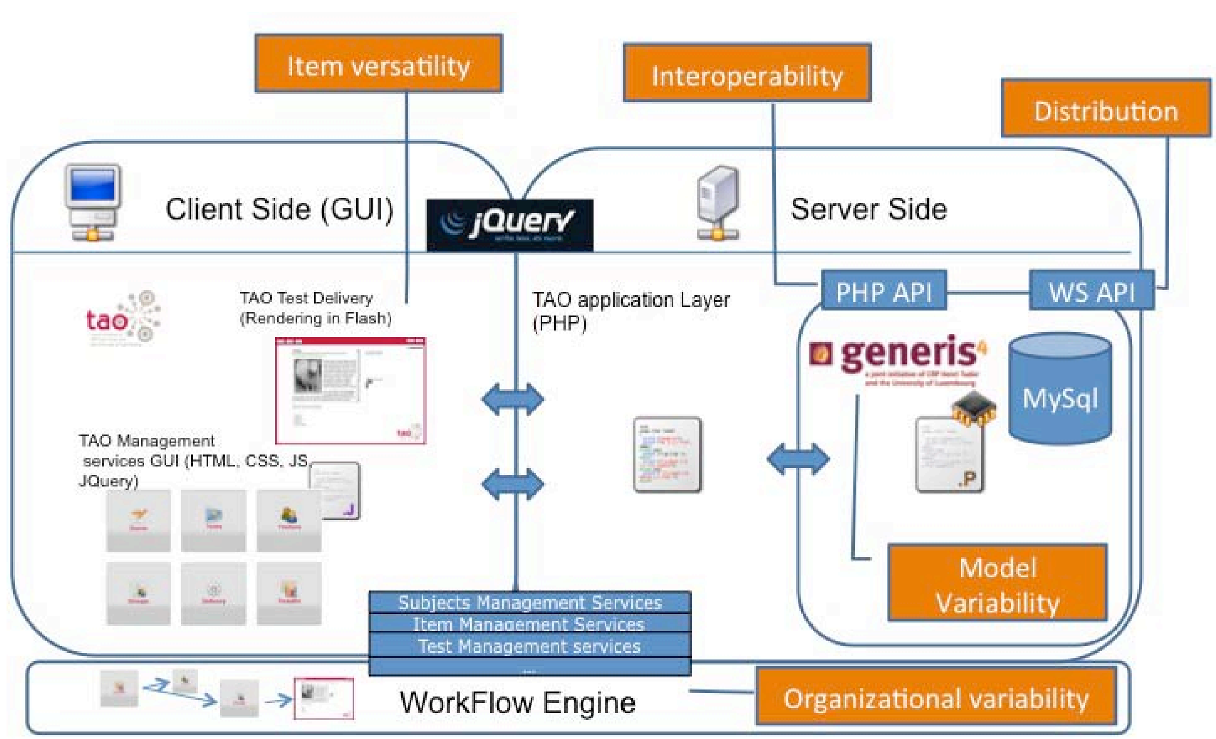
- Tackling technology-based assessment challenges with the semantic Web
- A workflow engine for cross-cultural large-scale assessment
- Test and item delivery architecture

9.2.1 Tackling CBA challenges with the semantic Web

Considering the range of variability of CBAs, a versatile architecture that makes use of specific innovative technologies was designed to tackle the needs of the different stakeholders.

Several types of variability need to be addressed, not only in the context of assessment, but also to support the integration of CBA in e-learning environments (see Figure 9.2):

Figure 9.2: CBA challenges in terms of variability



Context variability: In the educational context, for instance, many different data models are created and managed such as the specification of the classrooms where the students sit, their teachers, their learning options, and so on. In addition, modern learning environments provide personalized learning situations that enforce the need to design assessment items accordingly. The annotation of the item could be used to personalize the assessment instrument, such as item selection or item layout adaptations that reflect, for example, disabilities or learning styles. These annotations could also be used to select more appropriate learning activities based on a student's performance on the different items. This model variability also applies to other resources in a CBA system such as subjects, tests, or management of the test results. From an IT perspective, this model variability is challenging because it prevents the definition of the data models a priori and the design of a classical database structure. To tackle this variability, we used the semantic Web-related technologies RDF and RDFS (which stand for Resource Description Framework Schema.) Both are standardized languages that enable us to express information about resources at any level of abstraction. They allow the system users to define the data model (i.e., to define classes of resources and describe their properties) as well as the data itself (e.g., to define values of properties that describe a particular student). All these data models are defined by using RDF/RDFS. They are created and adapted easily by the users of the system through intuitive user interfaces – no further implementation efforts are needed when the data model requires changes.

Using RDF repositories instead of a classical database design solves the model variability issue. The TAO platform makes use of the generis4 RDF/RDFS repository. This implies that, from the point of view of the application layer, the source code needs to be independent of the model, and all the user interfaces for resource management need to be generated by first inspecting the model that was defined by the user.

Interoperability: CBA systems need to be integrated into existing business processes and legacy software. This may also require replacing some existing feature subsets or extending the system with new features, which involves defining Application Programming Interfaces (APIs) within the architecture that provide customized plug-ins to handle all the CBA resources (e.g., for computing statistics, creating proficiency reports, using existing subject databases, etc.). TAO provides two APIs, one for direct calls with the PHP programming language and one for remote access using SOAP Web services.

Distribution: CBA involves pedagogues, psychometricians, statisticians, item encoders and item translators, and may also involve stakeholders located in different geographic areas, such as a ministry of education guaranteeing access to subjects, a pedagogic institute defining tests, private software companies creating rich media for items, and so on. All the different stakeholders manage different kinds of resources. This requires all the CBA resources to be distributed on the CBA platform across a network of different collaborating institutions. Subject management, for example, is probably allocated to a specific accredited institution. Item management may be located at a different site to prevent items from being stolen. Such a distribution of tasks calls for the architecture of a CBA platform to be modular and distributable using the existing communication channels. This can be tackled through the use of Web services. Such distribution also allows sharing of resources and can be combined with a peer-to-peer network protocol, which would enable, for example, test creators to search across the entire network for items based on item model properties.

Organizational variability: The involvement of different stakeholders may require workflow-based work in order to make sure that the right person accesses the right feature of the CBA platform at the right moment. This also addresses the need for a quality layer to optimize the processes that lead to the creation of a measurement instrument. From an IT perspective, this involves the use of a workflow engine tool and a process design tool so that the person responsible for the assessment can design the CBA process according to his or her needs. This also requires that features from the CBA platform be split into autonomous services that can be triggered independently.

Item versatility: Authoring of items should not be restricted by the specificities of the CBA platform. It should allow simple item creation, like multiple-choice questions, but also more complex items, such as simulations. Maximum freedom should be given to the item developers at the level of the item layout and structure, as well as at the level of item behavior (interactivity). This can be achieved by a) defining a high-level language that supports layout, structure and behavior description, and b) implementing an authoring tool that facilitates the design of items as well as an interpreter capable of rendering such items.

9.2.2 A workflow engine for cross-cultural large-scale assessment

9.2.2.1 Introduction

With the advent of large-scale, cross-cultural surveys (e.g., IALS), Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS), PISA and PIAAC), CBA becomes more strategic than ever before. While survey designers are used to facing adaptation and translation issues with paper and pencil, introducing the computer for such activities brings up new challenges.

Ideally, the very same version of a computer-based educational survey would be carried out in many different countries by simply translating its components. Unfortunately it might not be that simple. At an international level, cultural distinctness among areas most often lead to heavy modifications on how data collection in educational assessment will take place. From deleting questions to inserting brand new sequences of items midway through the assessment of some respondents, altering the original design of the material might be a frequent need. Depending on local constraints, participating countries will sometime have to modify how questions are asked and sequenced. Software instruments must also deal with a wide range of alphabets and symbols and need to be highly polymorphic to satisfy all stakeholders collaborating on cross-cultural large-scale surveys.

Deriving as many instruments as necessary to cover all countries' specific needs (lingual, cultural, socioeconomic, etc.) can be painful and error-prone work if not managed properly. To transpose a paper-based instrument into a computer-based one is also expensive in the context of a large-scale survey. To minimize the investments in time and money, a good approach would be "describe once, adapt as needed, and run many." To reach this goal, we have adopted a solution based on *workflows*, easing the production of extendable and executable assessment processes. By both adapting and running BQ and CBA built on the same computerized descriptions, we have facilitated the design and implementation of cross-cultural computer-based educational assessment.

In the next sections, we describe how concepts and techniques from the world of workflow modeling and execution were applied to PIAAC. We introduce the formal representation of the assessment process based on the XML language. Then, we show how this format is used to support cultural adaptations. Finally we describe the execution of an assessment process in a computer-based context.

9.2.2.2 Describing the assessment process

Large-scale international computer-based surveys involve multidisciplinary teams from various countries across the world. This may lead to cultural and linguistic issues at the very beginning of the survey design; this adds up to the drawbacks that international projects usually encounter (e.g., agreeing on time-zone windows for meetings). The need for a formal representation about how data will be collected from an assessed population appears at an early stage of the engineering. A semantically sound pivot format giving stakeholders an opportunity to interact with a way marked assessment process description is essential.

According to Deelman, Gannon, Shields, and Taylor (2009), a workflow refers to defining the sequence of tasks needed to manage a business or computational science or engineering process. This also applies to BQ and CBA worlds by treating questionnaires and assessment sessions as sequences of tasks to be achieved by respondents. In this context, we say *assessment process* is the sequence of tasks required to assess and collect data for a single respondent. This statement is the basis for a simple but semantically rich format used to describe complex sequences of the tasks composing an assessment process and the rules that steer its flow.

A. Main concepts

Our XML description language integrates elements that authors can combine to describe processes including activities, transitions, consistency checks, variables and derivation rules. The following list explains the semantics of these concepts in detail and how they can be applied to the assessment world.

Process

It encapsulates the whole sequence of events occurring during the assessment process. It contains a set of activities logically linked by routing rules. It also holds the context of its process variables. The combination of the variable values and the current activity of a process instance constitute its state. In BQ and CBA contexts, a process represents an assessment session focusing on a single respondent.

Variables

They are comparable to simple data holders within a process. Our implementation gives every variable a global scope (i.e., variables are available within the entire process). Data collected during assessment will be stored in process variables.

Activity

It represents a task to be achieved during an assessment session. Activities are started according the flow logic of the process. They are divided into two categories: *Automated activities*, inherent to the process logic; do not need human interaction (Silver, 2005); they contain derivation rules or consistency checks. *Interactive activities* require human interaction and are focused on the respondent. For instance, such an activity could be an electronic reading item or a multiple-choice questionnaire.

Transition

It links activities using simple or complex logical rules. They might be used to direct the respondent to more or less difficult items during a single assessment process according to previously collected data, stored in variables.

Consistency check

It verifies the consistency of previously collected data. A violated consistency stops the assessment flow. Performing such checks at runtime is particularly useful for large assessment processes, including for the profiling of the respondent.

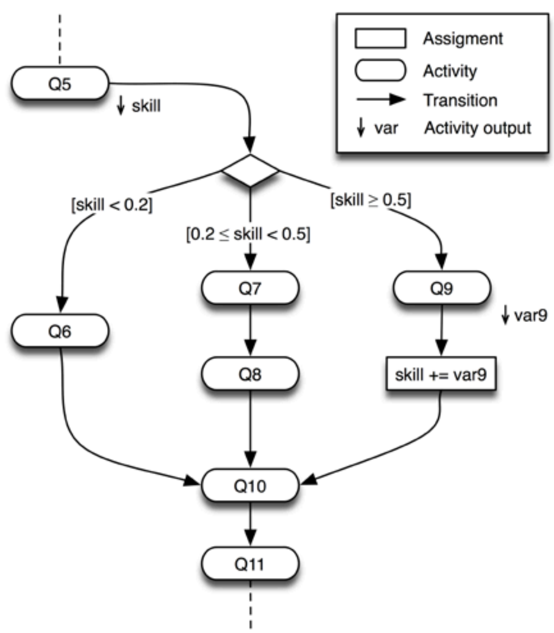
Derivation rules

Using derivation rules, authors can force the creation of variables at runtime or the assignation of a new value. Therefore new data can be generated based on values previously collected at assessment time. These derived variables will be used for subsequent derivations and flow control.

B. XML Description Language

Our XML syntax is used to depict assessment sessions. It is inspired by standards like BPML (Arkin, 2002) and XPDL (2002) that are currently developed and used by the industry sector.

Figure 9.3: A simple assessment process representation



Specially, the XML Process Definition Language (XPDL) specification offers a simple and minimal set of constructs present in most workflow products (van der Aalst, 2003). We used the following XPDL elements in addition to new ones to describe assessment processes (see Figure 9.3):

Workflow process

It contains the definition of activities, transitions and assignments.

Activity

It represents interactive steps and refers to the relevant software piece on which the respondent will interact.

Transition

It binds activities, transitions, assignments and consistency checks together using conditional expressions involving previously collected variable values.

Assignment

It contains logic setting a value to a particular variable by involving previously collected data. This is comparable to derivation rules but using XPDL semantics.

Consistency check

It verifies the consistency of the data contained in multiple variables.

Variable

It contains data collected during the assessment process. Transitions, assignments and consistency checks rely on their value to build up their logic.

9.2.2.3 Authoring and cultural adaptation support

To meet cultural requirements, stakeholders in large-scale and cross-cultural surveys need the right tools for cultural adaptations. They would like to add new cognitive items to the core survey in a particular sequence or suppress optional parts. Other partners may also need to add profiling questions related to the historical background of their country. All these adaptation needs are very difficult to foresee at the early stages of the assessment design. In this context, we developed appropriate tools to author, adapt and review XML-based assessment process descriptions.

A, Authoring companion tools

For the design and adaptation of assessment processes, we provided an authoring tool that did not require mastering the underlying XML format. Accessible via a Web-based user interface, it guides the author through the authoring process. Authors can comment on every modification they performed on the process description. This acts as a track-changes system, improving the communication between stakeholders. The authoring tool also prevented logical errors from occurring. For instance, a reference to an unknown variable in a flow transition was systematically detected and made visible.

Because activities, transitions and variable assignments may be numerous in a large assessment process, it was difficult to display all at once. To compensate for this shortcoming, the tool generates a graphical representation of the process based on the GraphML language (2010). This offered an easy way for authors to visualize the relationships between very distant elements of the process and to understand immediately its global design. This XML-based graphical representation is seamlessly produced using the XSLT technology (Clarck, 1999) aiming at transforming XML files using style sheets.

B. Continuous control

Creating numerous adapted versions of a reference assessment process (called *master*) must be carefully controlled. Critical aspects such as the comparability of collected data have to be taken into account. Thus, frequent reports were produced to set up a step-by-step reviewing process. These reports helped reviewers to validate or not the changes made by adapters. To support this, a reporting tool able to detect differences between multiple versions of an assessment process description was created. To produce its reports, the reporting tool uses flat and hierarchical

Diffing techniques (Miller & Myers, 1985; Chawathe, Rajaraman, Garcia-Molina, & Widom, 1996) in addition to the Levenshtein (1965) algorithm. As a result, additions, suppressions and modifications performed on an assessment process description are consistently underlined. Finally, the difference reports are delivered to reviewers in an XHTML format (Pemberton et al., 2000), usable on the computer but also easily printable.

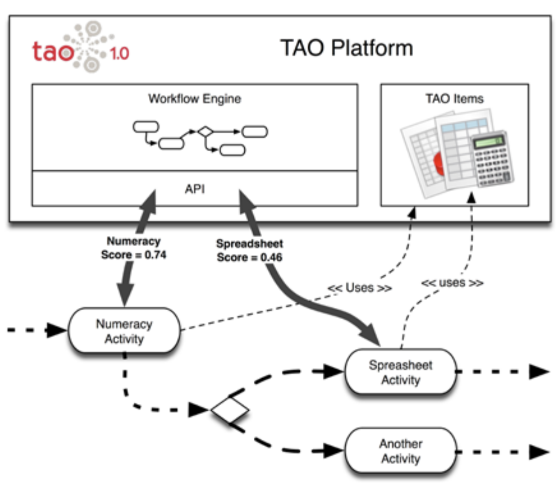
9.2.2.4 Assessment process execution

On top of our workflow description format, we built up a specific mono-user workflow engine (WE). It acts as a software service providing the runtime execution environment for a process instance (Ferreira & Ferreira, 2004, pp. 309-324) focusing on the assessment of a single respondent at a time. Formerly described by authors, the assessment process is executed exactly as is. This helps to reduce the gap that may appear in computer-based projects between project designers and computer scientists.

A. The workflow engine

Our workflow engine is developed upon the PHP programming language (2010a). It was designed to be reusable as a Web or desktop (e.g., in a software built upon PHP-GTK) (2010b) application component as well. As a consequence, it can be embedded on both server and client sides. XML descriptions of the assessment process are processed and stored in the system, ready to be instantiated and run. As a test bed, we integrated our component into TAO: an open and versatile CBA platform (Latour & Martin, 2007, pp. 32-33). This enables users of this PHP platform to author workflows to drive simple to complex assessment processes with pre-existing TAO items as its interactive steps. As with any WE, it is in charge of instantiating assessment processes on demand. Each process represents the test session of a single respondent. These instances can, of course, be started, paused, resumed and destroyed at will. Each process has its own state, composed by its process variables values and its current activity. The engine takes care of sequencing activities in the correct order according to the process state and the routing rules it comes across (see Figure 9.4).

Figure 9.4: The workflow engine in the TAO platform



Activities depicting the tasks that have to be achieved at runtime are executed differently following their nature. A dedicated "rule engine" embedded in the software executes the logic bound to automated steps and evaluate routings. On the other hand, the execution of interactive steps is delegated to the appropriate software component. Currently, the engine only supports Web applications for "interactive steps" implementations. They will thus be provided to the respondent through a Web browser during the assessment process.

At runtime, rules attached in automated steps have an inherent access to process instance variables because of their technical proximity with the WE. By contrast, interactive steps are software parts developed apart from the engine. They need easy access to process contexts in order to influence the process flow if needed. To make this possible, the WE exposes an appropriate API. Through the latter, interactive steps implementations are able to consult and act on the process state by retrieving and modifying the value of its variables.

B. Assessment delivery

The assessment is delivered to the respondent through a Web browser. The graphical user interface (GUI) displayed to the respondent during an assessment session is the process browser (PB). It is in charge of proposing the correct interactive steps (e.g., cognitive items, questionnaires) to the respondent. The PB enables the user to interact with interactive steps implementations but also provide GUI components to navigate through the process. This allows the respondent to go back and forth between interactive steps and adjust previous answers if necessary. Each time the respondent decides to go forward through the assessment using the GUI, the WE gets the control back and selects the next interactive steps to be displayed.

9.2.2.5 Conclusion

We have reviewed the methods and techniques coming from the domain of computer-based workflows and how they were successfully applied to the PIAAC survey. The XML format and its semantics for describing assessment processes were satisfactory. Tools provided for the authoring and the cross-cultural adaptations of PIAAC processes were adequate. Finally, the workflow engine provided the expected support in that context.

9.2.3 Test and item delivery architecture

It is mandatory for tests and related items to be delivered the same way on any Web browser in order to prevent biases and discrimination. It is also important to be able to track the respondent's behaviors (for CBA) or hesitations and latencies (for BQ) as those data may cast some light on subsequent results' analysis. Moreover, to give a maximum of freedom to the authors for the design of rich items including media content is very desirable. Those constraints have led to the choice of the Adobe Flash technology to create and deliver the test to the end user. The test and the item runtime engine read XML descriptions of the items that are authored beforehand with authoring tools. The XML-based items are based on the Business, Layout, Actions, Content, Knowledge Model (BLACK), which was developed at CRP Henri Tudor. The BLACK model is a high-level pivot format used to address the need for freedom required at the level of item creation. Thanks to the authoring tool, the user can edit items graphically and does not have to handle the item description file directly.

BLACK is composed of different sections. The *business* section of an item description gives overall information about the item. It describes those parts of the item that constitute the stimulus (set of materials given to the subject, the part of the item that constitutes the task itself, the response categories). In addition, it also defines the right expected answer and the evaluation algorithm to be used. The *layout* section describes all graphical elements to be used in the item such as radio buttons, check boxes, images, and so on by using the XUL markup language

(2010). The *action* section is used for items representing simulations, where certain elements of the item are expected to have a certain type of behavior. For instance, a pressed button might cause a Montgolfier picture to move up and down. The *content* section is language dependent and contains all text messages available in the item or all links to media, such as pictures or movies. Finally, the *knowledge* section contains metadata annotations of the item using XML RDF and describing potential skills, context of use, overall difficulty of the item, and so on.

9.2.4 Use of the TAO platform for PIAAC

9.2.4.1 BLACK model to support rich CBA items (component Problem Solving)

Motivations in using the BLACK format

The BLACK format addresses particular requirements of the PIAAC CBA platform to build and to run cognitive items, especially in the field of "Problem Solving":

- Dynamically interoperate heterogeneous contents: e.g., for complex objects assembly;
- Facilitate contents extraction and maintainability: e.g., for localization (XLIFF);
- XML driven – based on namespaces to favor natural plugin construct and interaction;
- Encapsulation to assure packaging, transportability and deployment of the resources;
- Hierarchical nesting of contents from various data sources to favor interchangeability;
- Events-driven to assure the interoperability between the components;
- Homogeneous deployment on server side as well as on the client side.

BLACK — a MVC architecture

BLACK is an extension of the MVC pattern; it respects and elaborates on the MVC architecture principles. From a general point of view, the layers Business, Layout, Action, Content, and Knowledge can be successfully mapped to the Model, View, and Controller layers, as defined in Table 9.1.

Table 9.1: The mapping between BLACK and MVC

Model	View	Controller
Business		
	Layout	
		Action
Content		
Knowledge		

BLACK data structure

Streams of data structured according the BLACK format are named BLACK formatted streams or simply BLACK streams. The content of a BLACK data structure is a *Manifest* which references zero, one or many layers of the BLACK model: Business, Layout, Action, Content, and Knowledge.

```
<?xml version=“1.0” encoding=“UTF-8”?>
<black:manifest xmlns:black=“http://www.exulis.lu/black.rdfs#” id=“an_example”>
 <black:business>
 ...
 </black:business>
 <black:layout>
 ...
 </black:layout>
 <black:action>
 ...
 </black:action>
 <black:content>
 ...
 </black:content>
 <black:knowledge>
 ...
 </black:knowledge>
 </black:manifest>
```

This structure constitutes a normal basic BLACK stream; the root tag (i.e., “manifest”) must expose at least two attributes: a reference to the BLACK namespace and the “id” of the manifest. In the following section, we briefly present the definition of a manifest’s prime layers and show how this supports our goals and assertions regarding the authoring and rendering of Problem Solving items described as BLACK bundles (i.e. one or more linked BLACK manifests and the relate resources (e.g. pictures).

The *Business* layer introduces the “semanticity” of all elements referenced by a manifest. Therefore, it operates as the driver of the integration and the interoperability of all the different kinds of contents involved in a bundle. It references the namespaces of the tags present in the BLACK stream and refers the components (e.g. parsers, services, etc.) to be loaded and invoked dynamically when certain namespaces need a specific handling (e.g. for the rendering of a graphical element.) The layer also defines the bindings existing between elements of *Layout* layer and data providers defined in the *Action* layer (e.g. services, functions) or existing as data container in the *Content* layer. This *Business* layer will also hold all the preferences and settings subject to adaptation by the user and/or the system. It also contains directives for the correct handling of the BLACK stream itself; the directives indicates, amongst others, where the BLACK stream should be handled: on the server side or on the client side. When a BLACK stream is processed on the server side, a specific parser, named *BLACKparser*, analyzes the stream and prepares a deployment infrastructure without any reference to the BLACK format so the final product results in a standard web application where the use of BLACK becomes completely transparent to the server and the client. The BLACKparser is thus in charge to go from a meaningful compact BLACK manifest to a multi-tier application, and therefore, adding all the necessary in-between components enabling the expected execution of the whole application for the final user. A BLACK stream can also be carried along to the client side. In that case, the BLACK stream is handled by a *BLACKrunner* module. The BLACKrunner is available as two different flavors: one is based on the Adobe Flash technology and the other is built on a JavaScript technology. The Flash version of the BLACKrunner embarks the eXULiS library which natively handles the XUL and the SVG formats (even on browsers not compliant with these standards) and offers an extension mechanisms (to render other formats) and a toolbox facilitating rapid duplex communication, local storage, data processing, and enforcing resilience and security. The JavaScript version of the BLACKrunner is more dependent on the web browser capabilities (e.g. the availability of the canvas object) but it clearly offers a broader compliance to the W3C standards on which it tries to capitalize a maximum (e.g. XHTML, CSS3, HTML5, and so on.)

The *Layout* layer is dedicated to the declaration of all the visual components displayed to the user (i.e. the side which was referenced as a part of the "Editor" in the first edition of the MVC (Reenskaug, December 1979) and was re-centered on the term of "View" in the second version of the MVC (Reenskaug, May 1979). These elements are usually related to a part or a complete framework offering graphical rendering capabilities; frameworks commonly available are XUL, SVG, XAML, XHTML+CSS, MXML, etc. Some of these frameworks are dedicated to specific domains as MathML, ChemML, and ChartML, in which they may need the support of an extra plugin in order to insure the right display of the elements. The rendering of very advanced standards as the X3D and VRML also require the availability of specialized renderers installed on the client computers. When a BLACK stream is used as a support for a mashup aggregating various sources of contents, the Layout layer may summarize itself to a set of XHTML iframes (each pointing to its own contents) placed accordingly to a CSS. For this case, the added value of the BLACK is to enrich the default behaviors of each iframe by the injection of an *observer pattern* making the content of each iframe potentially reactive to the interactions occurring in other iframes (i.e. on other source of data.) An illustrating use case could be this one: let's suppose English is not your mother tongue. You connect to a portal offering a mashup specially designed to facilitate the reading of scientific texts available in English. After a light setup where you define your linguistic preferences, the reading facility would display two iframes; one is displaying a *Scientific American* article, and a second targeting *Google translate* online application. Each time you would underline one word of the article, an event would be dispatched from this first frame; the event would be broadcasted by the BLACKrunner; an event listener attached to the second iframe would react and, as a result, the application Google translate would automatically display the translation of the highlighted word in your preferred language.

The *Action* layer is the container of all the logic processing. As the BLACK elements must be as loosely coupled as possible, all the interactions between the elements are assured by a proper events manager which is built on event dispatchers, event listeners, and if necessary event broadcasters. The logic (e.g. function, service, object method) registered to handle the catch of an event by a listener is naturally located in the Action layer. Also, all forms of processing, controllers, data access, etc. are nested in this layer. The BLACK format defines a limited set of elements which stands as a meta-language used to specify context of execution (i.e. server side or client side), loops, alternatives and control structures (e.g. “if/then/else”, “switch/case/default”), standard service calls, and so on. However, application logic needs more powerful capabilities. Therefore, it is possible to embark logic expressed in any language in a CDATA structure as long as it does not break the XML validity of the BLACK stream. Thus, because of its capacity to provide data via functional and service calls, the Action layer does not restrict to a pure Controller tier of the MVC architecture, but could also be considered as a part of the Model. Nevertheless, in our view, we prefer to assimilate all forms of processing, data manipulation, and service invocation present in the Action layer to the Controller part of the MVC architecture.

The *Content* layer is another part of the Model tier of the MVC architecture. Within the BLACK format it acts as a pure data container or as set of references to externalized data which themselves can be other BLACK streams. This may lead to extremely complex bundles made of imbricated BLACK manifests. In use cases where contents must be adapted depending on cultural constraints, it became totally obvious that all the data prone to localization and translation had to be located in this layer to facilitate their extraction by specialized processes (as met, for example, in international large-scale surveys) and their subsequent re-injection.

The *Knowledge* layer is a container for all the metadata describing the BLACK stream as a whole or parts of this stream (e.g. a description of a picture referenced in the Content layer, displayed in the Layout layer and voiced by a call – declared in the Action layer, to an external text-to-speech plugin announced, via the Business layer, as a capability to be initialized for the current BLACK stream.) This layer also specify the format of the data to be produced as an output (if any) of the system described by the BLACK stream, for example, events to be collected.

9.2.4.2 eXULiS – a Rich Internet Application (RIA) framework used eTesting

Rich Internet Applications (RIAs) seem to be the solution offered by computer-scientists to the more and more demanding users. Unlike previous generations of Web applications, RIAs plan the leverage of the user experience with more powerful graphical user interfaces (GUI) including charts, drag and drop. RIAs also aim to remain as platform-independent and setup-free as possible. One should be able to work on data and tools both available online and these tools must be usable without setup or configuration on the client computer. However, RIAs also should be more than a few good-looking GUI because the desktop applications have made the users accustomed to a high level of responsiveness and customizability.

RIA solutions are now legion. These are proposed to the developers as frameworks. Nearly none of these frameworks is really standardized even if the majority relies on the XML as a base for at least the GUI mark-up. The well-known application/UI formats are Mozilla’s XUL, Microsoft’s XAML, Adobe’s MXML, Laszlo Systems’ LZX, ActionStep’s ASML, ASWing’s AWML (used for PISA 2009) and enFlash’s ML.

For the Problem Solving component of PIAAC, we used a new framework named eXULiS. It was not a standard or the ultimate RIA solution but rather another alternative that had been matured since 2006 and that tried to stay as close as possible to the standards of the moment. This section gives an overview of eXULiS and shows its place in the PIAAC assessment.

The overall architecture of the TAO platform

The Research Centre Henri Tudor started to develop a CBA platform called TAO (in French "*Testing Assisté par Ordinateur*") since 2002. Until 2011, the platform relied on two main components. The first part, responsible for structuring, storing, and sharing of the data, is located on the server side. It is called Generis.[4] The second part, active on the client side, was in charge of providing the correct display and completion of the tests and the questions (or items). It relied on the Flash player plugin that is installed on a large computer base (amongst individual computers). Since late 2011, more standard HTML5+JavaScript+CSS technologies replaced this latest component.

This section describes client-side technologies as used for PIAAC, prior to December 2011.

The work of the client-side component is to receive and to interpret the files describing the tests. The syntax of a test description is a mix of XUL and QML. XUL (or "XML User Interface Language") gives the layout of the graphical objects to appear on the screen of the computer for the test. We choose XUL for its maturity and because it was more open and community oriented than some other initiatives. QML stands for "Questions Markup Language." The QML used in the TAO platform was extended to encompass the needs of the IRT model as well as the new requirements towards the multilingual assessment.

The heart of the rendering engine was a parser called XUL2SWF (where SWF is the file extension of the Flash movies). The framework eXULiS is more than a simple evolution of the XUL2SWF engine. It contains a XUL parser completely rewritten to be extensible and more compliant with the Mozilla specifications. It also includes a second parser that is able to display SVG drawings. SVG is the acronym for "Scalable Vector Graphics." This XML language is used for describing two-dimensional vector graphics. The integration of a drawing format was required to open new vistas for the design of RIAs and for the authoring of advanced types of tests and items in the PIAAC context.

The authors who create new graphical layouts for RIA and/or for CBA can proceed using the tools freely available on the Web; XUL files can be written with xuledit.xul and SVG files can easily be produced with InkScape, for example. For PIAAC, we also developed a tool to adapt the XUL and SVG layout files. This tool was named Copernicus.

Figure 9.5 "TAO platform architecture" depicts the interactions of the two components described here above in order to deliver the RIAs, and more specifically, the tests to the test-takers. This schema shows the topography (or deployment architecture) and some sequences of actions between the modules involved in the TAO platform. This schema should be read from left to right as everything in our platform starts from Generis (Plichart, Jadoul, Latour, & Vandenabeele, 2004).

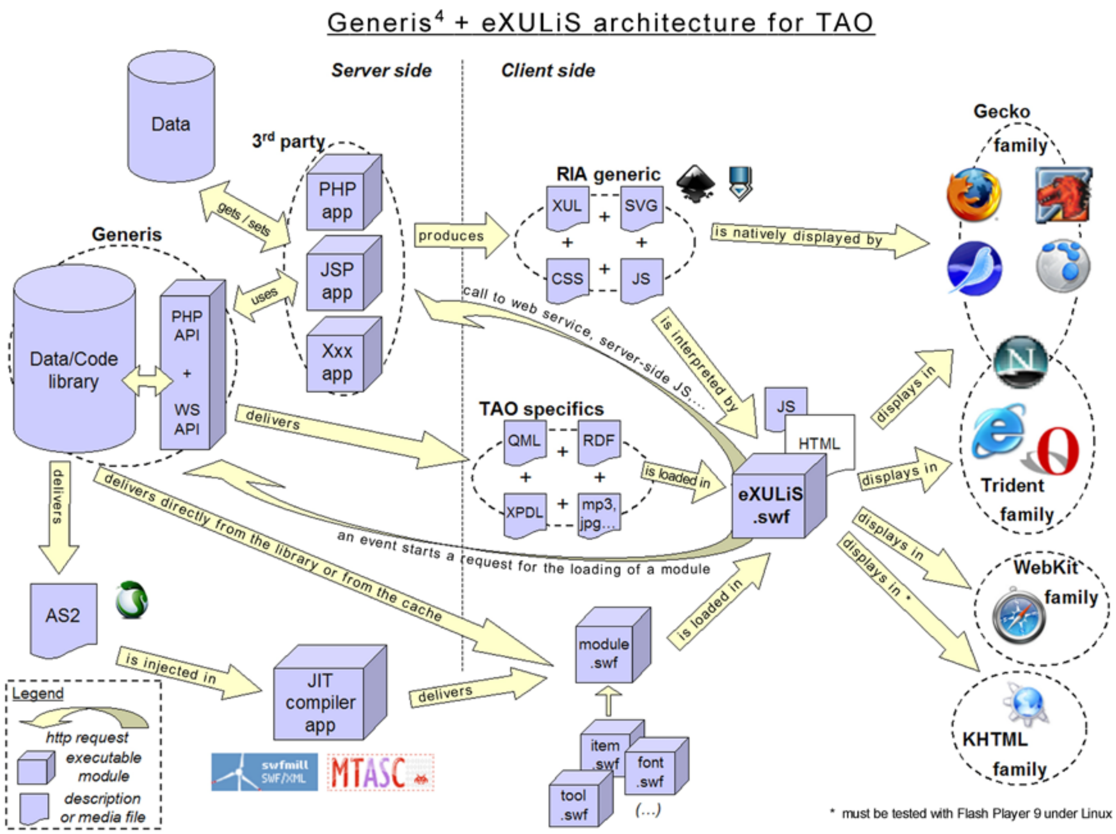
On the server side, Generis provides, via its PHP API or via its Web Services API, two sets of information issued from the RDF triples it manages. The first of these sets of data can be used by 3rd party applications (Authoring tools, eLearning platform, eBusiness applications, and so on) to produce the RIA description files (XUL, SVG, Cascading Style Sheets and JavaScript) that can be either natively rendered by the browsers embarking the Gecko Runtime Engine technology or that can be displayed in all the families of browsers via the eXULiS plugin. The second set of data contains the same kind of files as the ones involved in the first set but it also includes files delivered in a format that is not directly interpreted by the Internet browsers (even the Gecko family). This part of the second data set is solely dedicated to the eTesting. It holds some definitions that only have a meaning in the context of a CBA.

The files formats specifically used by eXULiS for CBA are: the TAO QML definitions, the XPDL (XML Process Definition Language) and some specific XML and RDF datasets. Please note that, in the case of PIAAC, RDF information as well as XUL and SVG description are conveyed by the BLACK files described earlier in this document.

TAO QML files contain the logic and the hierarchical structure of the assessments. It means that the files describe a specific assessment in terms of a campaign involving one or several sequences of tests (potentially in different languages) including one or more sequences of Items made of a set of particular Items, each one integrating a Problem (stimulus), and Inquiries composed of a Question and a Distracter (e.g. a set of Proposals for multiple choices question, an open text, a puzzle, and so on). TAO QML is described later in this document

In the case of predefined sequences (called scenarii) of Tests, Items or Inquiries, eXULiS evaluates, at a moment T, the execution context of the assessment; then it uses some Workflow (WF) definition files formatted in XPDL that contains the conditions of the time T, to display the correct user interface at time T+1. As mentioned above, the definition of the GUI is not stored in XPDL but in XUL and SVG files.

Figure 9.5: TAO platform architecture



Another interesting aspect of eXULiS is its capacity to extend its dynamic behavior in numerous manners. The engine is a Flash movie (.swf file) and it can act as a relay to local or remote function calls. It embarks a set of wrappers and API that enables the invocation of Web Services, server-side JavaScript, local JavaScript (located in the Web page – HTML, PHP, JSP – that nests the eXULiS Flash object), remote CGI scripts, and it even allows the communication with client-side desktop applications through the use of local connections.

When it detects a specific need (for example, a Test event is raised requesting that the current Item is displayed in Japanese), eXULiS may forward this event to the Generis back office that will provide (if necessary via a Just In Time compilation) the useful resources (in this case a *.swf* module containing some Hiragana, Katakana and Kanji character sets). The .swf modules may contain diverse types of resources including fonts, tools (e.g. calculator, notepad), compressed XML datasets, media, and so on.

In the next sections, we will briefly discuss the eXULiS and explain how its modular internal architecture is a favorable ground to extend its capabilities.

Design of the eXULiS framework

<u>A two-fold construction</u>

The effort to build the eXULiS framework started before the diverse initiatives led by the standardization agencies (e.g., W3C). Instead of creating a homemade GUI format for the specific needs of the CBA, we decided to select one that was already available. We wanted an open standard, well established with enough resources available online, and if possible, a solid community; the Mozilla project called XUL, although not a standard, got selected.

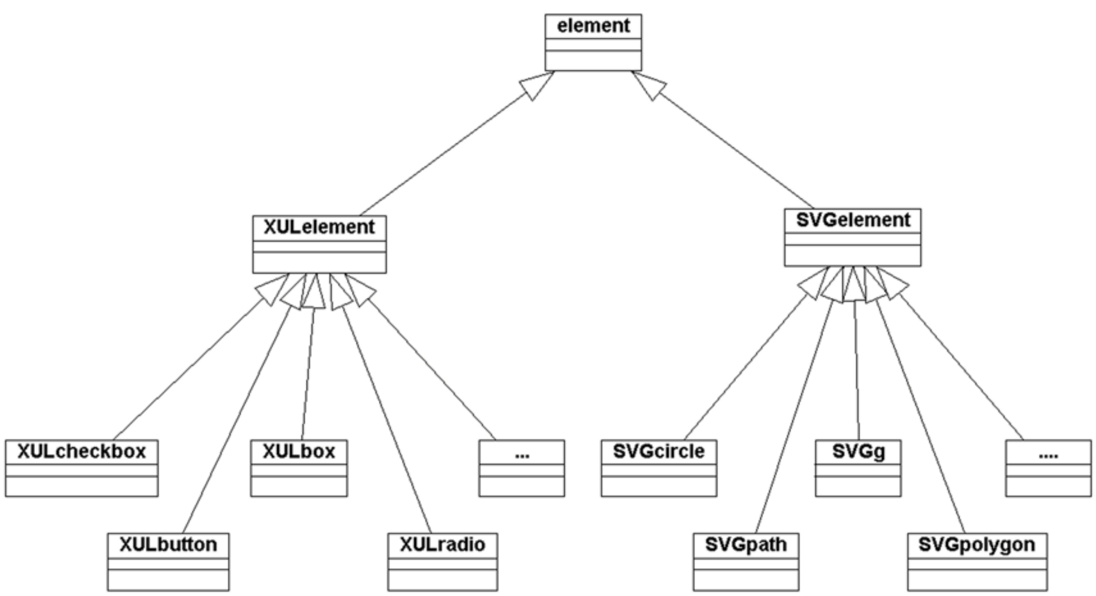
For our needs, we implemented this format in a component named XUL2SWF. Initially, this was a monolithic class encapsulating a parser that was recursively parsing XUL tags to render the corresponding graphical user interface via the Adobe (formerly Macromedia) graphical widgets (e.g. basic "movie clips" and "v.2 components"). In 2007, about 60% of the XUL widgets were available in XUL2SWF, already enough to fulfill most of the needs of PIAAC.

In 2008, the one-piece class became the eXULiS framework and the coverage of the XUL specification increases as well as the new engine embarks other XUL affiliated technologies like XBL (eXtensible Bindings Language) and RDF (Resource Description Framework).

Furthermore, to address the new requirements elicited during the PIAAC survey preliminary analysis, we added to the XUL framework, a second framework to handle the SVG standard.

In the Figure 9.6 "eXULiS framework overview," the two parts of the class tree can be clearly identified: on the left side, XUL classes are inheriting from a common XULelement, and on the right side, SVG classes are inheriting from a common SVGelement; both XULelement and SVGelement are inheriting from element. The ancestor class element acts like a relay allowing a natural communication between the two frameworks, in particular via events. For example, widgets in the SVG framework can subscribe to any events of the type xyz and start to listen while a widget in the XUL framework can dispatch this type of events. The xyz event bubbles up to the root of the frameworks and it gets broadcasted to subscribing SVG widgets.

Figure 9.6: eXULiS framework overview

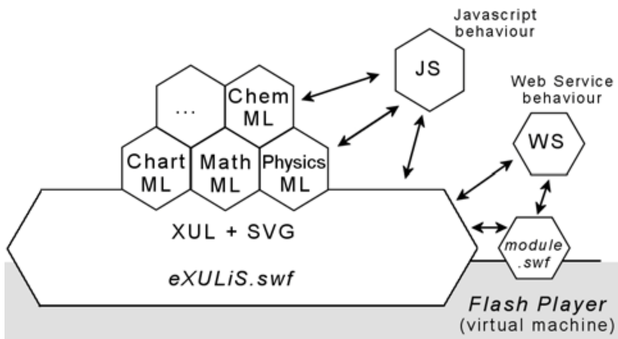


The whole framework can also use other static classes available in the Flash framework; an example is the Tween class that allows movements and transitions in the state of the widgets.

<u>How can eXULiS be extended?</u>

The integration of the SVG standard in eXULiS was first intended to allow authors to create their own custom themes and skins for the tests/items (e.g. a button with shape of a cloud for 6 years old children). The power of the SVG standard and the capacity of Flash to call some external JavaScript functions, unleashed the potentials of eXULiS. First, we created a module transforming some ChartML tags into SVG (see Figure 9.7) that eXULiS displays perfectly. Our latest projects target some needs in physics and genetics' laboratory simulations.

Figure 9.7: The eXULiS extension



An example of use of eXULiS

This example (on Figure 9.8) is a mix of physics rules (gravity, levers, axis) applied to a schema that is a composition of SVG drawings and XUL widgets (buttons, checkbox) interacting in a laboratory allowing to experience a problem of static physics science.

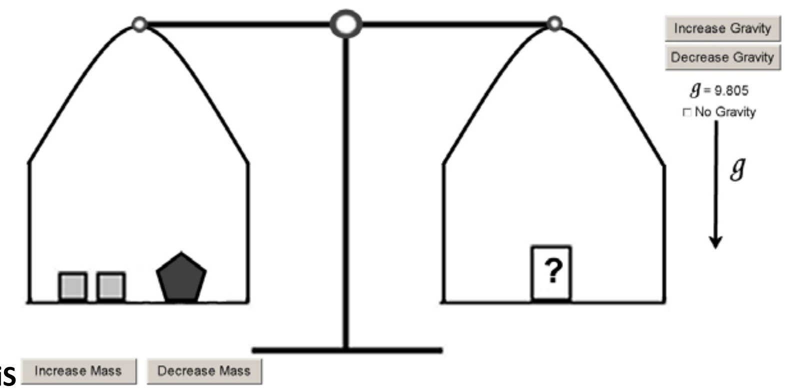
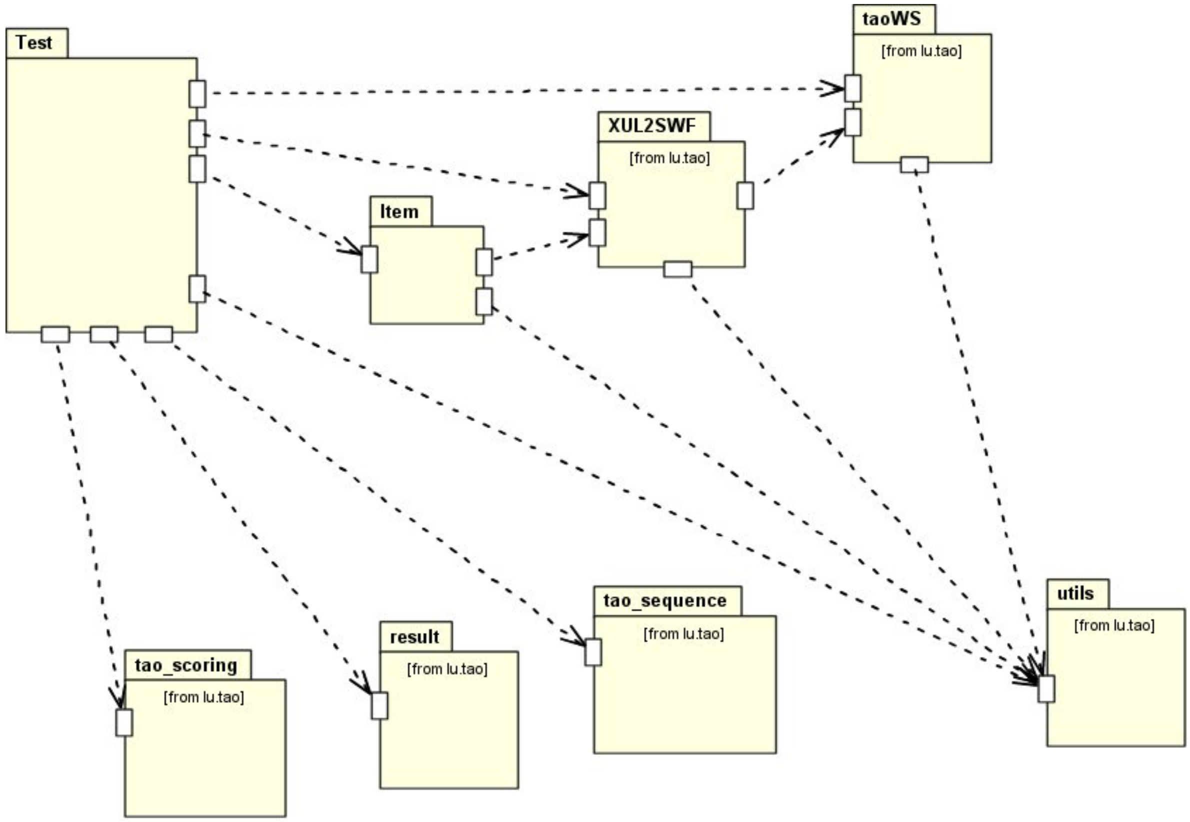


Figure 9.8: A physics lab with eXULiS

II.2.6Test & item execution engine

The packages' organization of the Test & Item Execution Engine (TIEE) is centered on two main packages, i.e. the Test and the Item (cf. Figure 9.9). The following description of the Item package is based on the architecture of the QCM (Multiple-Choice Question) item model. However, the Item package structure may vary with the design choices taken by the item models designers (e.g. if designers use XUL2SWF/eXULiS to create Problem Solving items). The only constraint for an item model is just to embark and connect its part of the communication interface in order to implement the communication protocol that must take place between the Test instance and any Item instances. This protocol is described in the figures hereafter.

Figure 9.9: Test & Item Execution Engine (TIEE) packages organization



The Communication Interface (based on Flash Local Connection mechanism) between a Test and any kind of Item (see Figure 9.10) is defined in an API (see Figure 9.11).

Figure 9.10: Test & Item Execution Engine (TIEE) packages organization

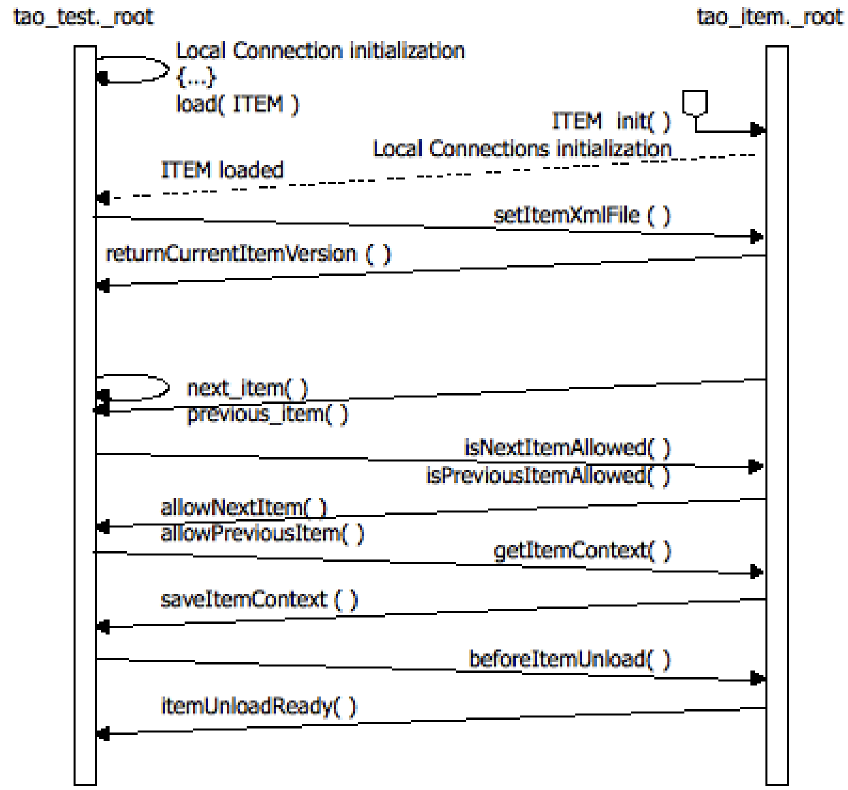
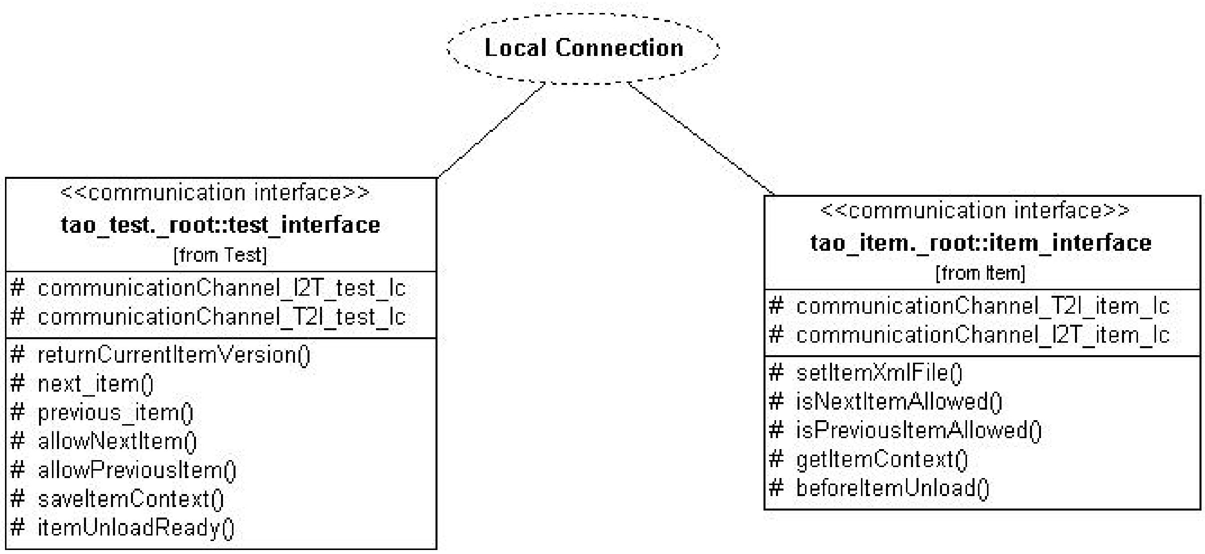


Figure 9.11: Communication API used in the Test & Item collaboration schema

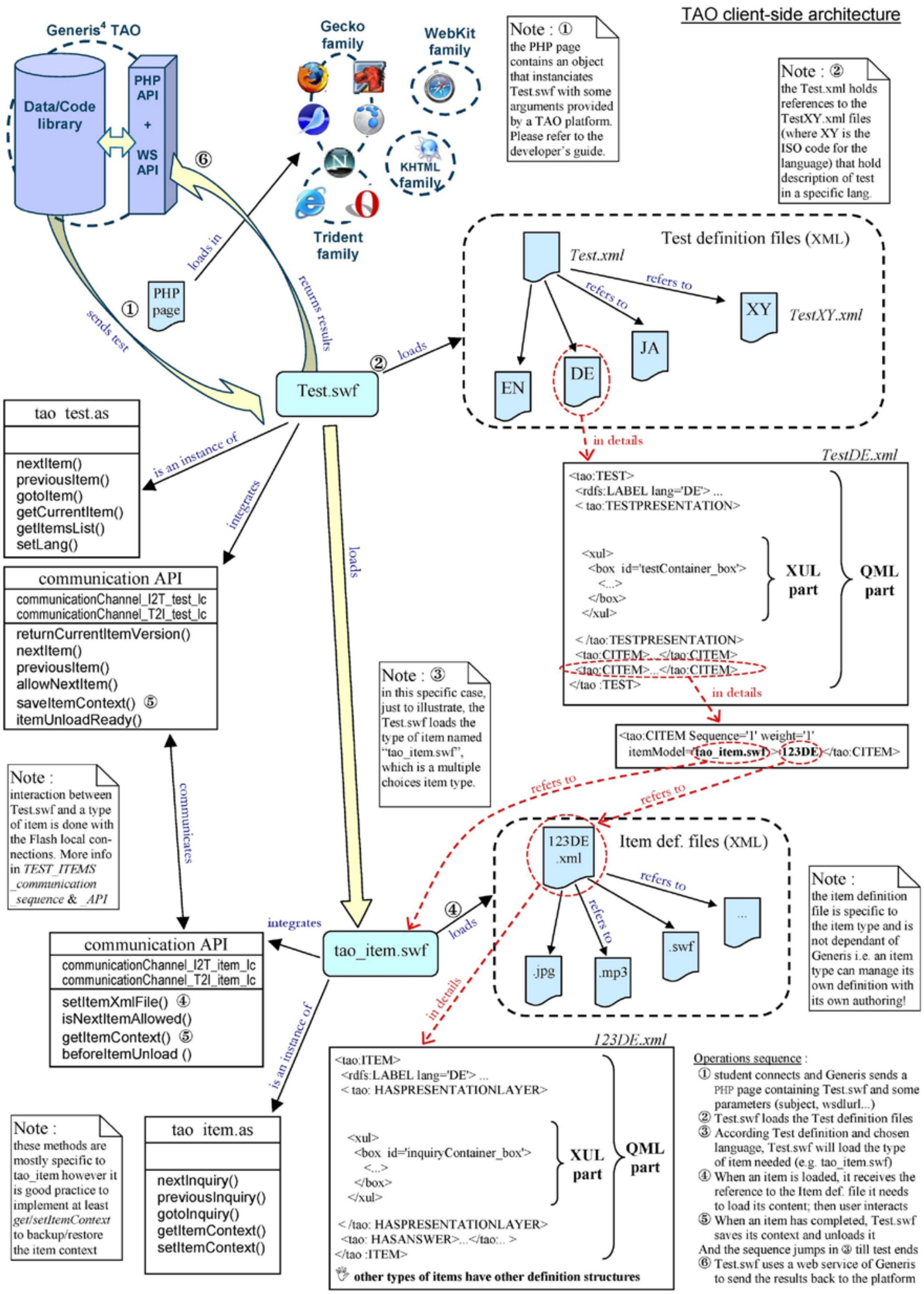


For the normal functioning of the TIEE, only the communication API (located in the item_interface, on the right part of Figure 9.12) shall be integrated in an item model. Thereby, some item models already available on the TAO platform, present slightly different architectures (e.g. the Text-with-gaps item model). However, a good practice is to separate the control process from the graphical rendering. It is also good to avoid to split the Item's side of the communication API or to merge it here and there in the different parts of the Item architecture. For PIAAC, the developers have applied Object-Oriented (OO) encapsulation paradigm, keeping the communication API in one layer of their item model movie clip, the GUI in a second layer, the data manipulation in a third one, and finally the control and process actions in one or more other layers.

In the case of PIAAC, although the Test is responsible for the activation of the scoring and items sequencing, the design of the TIEE is oriented to a delocalization of the scoring and sequencing algorithms in separated packages. This architectural choice is insuring a minimum maintenance cost while new scoring (largely dependent on cultural specificities; for example, the comma separator and the thousand separator vary from country to country; in Japan, even a ten-thousand grouping the numbers is usual) and sequencing methods are adapted or added to the TIEE.

The Figure 9.12 shows the standard execution and data flows of the testing process activities as managed by the TIEE. In the schema, at the level of the step 4, the structure of the item XML definition file (here, 123DE.xml) will vary depending on the design choices made by the developer of the item model; the model used in the example is the QCM (or Multiple-Choice Question) that uses the tao_item.swf Flash execution file.

Figure 9.12: Schematic illustration of the Test and Item execution



References

Arkin, A. (2002, 13 November). *Business process modeling language.* Retrieved from xml.coverpages.org/BPML-2002.pdf

Chawathe, S. S., Rajaraman, A., Garcia-Molina, H., & Widom, J. (1996). Change detection in hierarchically structured information. In J. Widom (Ed.), *Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data* (Montreal, Quebec, 4-6 June, 1996), 493-504. doi: http://doi.acm.org/10.1145/233269.233366

Clarck, J. “XSL Transformations (XSLT),” Version 1.0. (16 November, 1999). Retrieved from http://www.w3.org/TR/1999/RECxslt-19991116

Deelman, E., Gannon, D., Shields, M., & Taylor, I. (2009). Workflows and e-science: An overview of workflow system features and capabilities. *Future generation computer systems,* Vol. 25, 528-540.

GraphML Team, The. (2010). *GraphML specification*. Retrieved from http://graphml.graphdrawing.org/specification.html

Jadoul, R., Plichart P., Swietlik J., & Latour T. (2006). eXULiS – *A rich Internet application (RIA) framework used for eLearning and eTesting.”* Paper presented at the IV International Conference on Multimedia and Information and Communication Technologies in Education, m-ICTE 2006. 22-25 November, 2006. Seville, Spain; published in: Méndez-Vilas, A., Solano Martin, A., Mesa González, J., & Mesa González, J. A. (Eds.): *Current developments in technology-assisted education* (2006), Vol. 2. FORMATEX (http://www.formatex.org/micte2006/book2.htm), Badajoz, Spain, pp. 851-855.

Latour, T., & Martin, R. (2007, October). TAO, an open and versatile computer-based assessment platform based on semantic web technology. *ERCIM News, 71*, 32-33.

Levenshtein, V. (1965). Binary codes capable of correcting deletions, insertions, and reversals. *Doklady Akademii Nauk SSSR (Russian), 163*(4): 845-848. English translation in *Soviet Physics Doklady, 10*(8), 707-710.

Miller, W., & Myers, E. W. (1985, November). A file comparison program. *Software –Practice and Experience, 15*(11), pp. 1025-1040.

Pemberton, S., et al. (2000, 26 January). XHTML™ 1.0 The Extensible HyperText Markup Language, A Reformulation of HTML 4 in XML 1.0 (2nd ed.). Retrieved from http://www.w3.org/TR/2002/REC-xhtml1-20020801

PHP Group, The. (2010a). *Introduction to PHP*. Retrieved from http://www.php.net/manual/en/introduction.php

PHP Group, The. (2010b). *PHP-GTK: PHP-GTK 2 Manual: Preface.* Retrieved from http://gtk.php.net/manual/en/preface.php

Plichart, P., Jadoul, R., Latour, T., & Vandenabeele, L. (2004). Communication at the advances in intelligent systems – theory and application. *AISTA 2004.* Luxembourg.

Reenskaug, T. (1979, May). *THING-MODEL-VIEW-EDITOR - An example from a planning system*. *Technical note.* Xerox PARC. Retrieved from http://heim.ifi.uio.no/~trygver/mvc/index.html

Reenskaug, T. (1979, December). *MODELS - VIEWS - CONTROLLERS. Technical note.* Xerox PARC. Retrieved from http://heim.ifi.uio.no/~trygver/mvc/index.html

Silver, B. (2005, 17 August). *BPMS watch: Agility and BPMS architecture.* Retrieved from http://www.bpminstitute.org/articles/article/article/bpms-watch-agilityand-bpms-architecture.html

van der Aalst, W. M. P. (2003). Patterns and XPDL: A critical evaluation of the XML process definition language. The Netherlands: Department of Technology Management Eindhoven University of Technology.

Workflow Management Coalition, The. (2008). XML Process Definition Language, Version 2.1a. Retrieved from http://www.wfmc.org/Download-document/WFMC-TC-1025-Oct-10-08-A-Final-XPDL-2.1-Specification.html

XML User Interface Language (XUL) (2010). Retrieved from https://developer.mozilla.org/en/XUL

Chapter 10: Field Operations

Pat Montalvan and Michael Lemay, Westat

10.1 Overview

As with all aspects of PIAAC, countries were asked to comply with a set of Technical Standards and Guidelines (TSG) for survey operations/data collection. These standards can be found in Chapters 2, 8, 9 and 10 of the TSG. Part of the TSG included a quality assurance (QA) and quality control (QC) program for survey operations covering the collection of a range of information about the design and implementation of PIAAC data collection in each country via written reports, phone conferences and some in-person meetings. (Chapter 11 provides a detailed description of the QA and QC program which facilitated the collection of this information.)

This chapter presents information about the 25 countries/territories that completed the PIAAC Main Study data collection: Australia, Austria, Canada, Cyprus,[1] the Czech Republic, Denmark, England (UK) and Northern Ireland (UK),[2] Estonia, Finland, Flanders (Belgium), France, Germany, Ireland, Italy, Japan, Korea, the Netherlands, Norway, Poland, the Russian Federation,[3] Slovakia, Spain, Sweden, and the United States.[4] All the information presented in this chapter is based on data self-reported by countries as of 31 July 2013.

Sections 10.2 through 10.7 of this chapter provide highlights of findings with respect to data collection timeline, production and response rates, interviewer training, field management practices, staffing and supervision, nonresponse reduction, and fieldwork quality control. Furthermore, at the end of data collection, interviewers were debriefed on their PIAAC experience. This feedback is summarized in section 10.8. Finally, section 10.9 concludes the chapter with recommendations for future PIAAC cycles.

Furthermore, it is also important to note that there were deviations from the TSG with regards to data collection in most countries. Whenever deviations were identified by the Consortium, be it during the planning, training or implementation stages, countries were notified quickly via email or telephone conference or both. If possible, acceptable alternatives were identified; otherwise both the country and the OECD were notified of the potential problem. However, for the most part, key TSG guidelines or acceptable alternatives were followed by most countries.

---

[1] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[2] England and Northern Ireland are reported on separately at the request of the United Kingdom.

[3] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[4] Portugal and Chile, two countries which participated in the Round 1 Field Test, officially notified the OECD that they would not be conducting the Round 1 Main Study. They later rejoined PIAAC as part of Round 2.

10.2 Data collection timeline

Countries were expected to begin data collection on 1 August 2011 and complete fieldwork by 31 March 2012 (8 months or 243 days). Table 10.1 presents detailed information about each country's adherence to the data collection timeline.

Almost 60% of the countries completed the fieldwork by mid-April and the remainder by 24 November 2012. The actual length of the field period ranged from 79 days in France to 284 days in Sweden (average: 224 days).

The majority of countries did not start data collection on 1 August 2011 primarily because they believed that the vacation plans of many field staff and respondents would negatively impact production in this last month of summer. Seven countries (Austria, Estonia, Germany, Ireland, Poland, England (UK) and Northern Ireland (UK) began exactly on 1 August 2011. Four countries (Australia, Canada, Russian Federation[5] and Slovakia) began data collection in late fall for various reasons. Canada and Australia started in November and October, respectively, due to ongoing competing projects. Slovakia and the Russian Federation[6] began data collection in late October and late November, respectively, due to contractual and budgetary issues. France made the decision to begin data collection in September 2012.

Most countries concluded data collection by mid-April 2012. Nine countries ended data collection on or before 31 March 2012. Thirteen additional countries ended by 31 May, Sweden and Canada ended in June, and France ended in November 2012.

[5] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[6] See above footnote.

Table 10.1: Data Collection Timeline

	Fieldwork Dates		Duration (Days)
	Start	End	
Australia	1 Oct 2011	31 Mar 2012	182
Austria	1 Aug 2011	31 Mar 2012	243
Canada**	1 Nov 2011	30 June 2012	242
Cyprus[7]	1 Sept 2011	31 Mar 2012	212
Czech Republic	15 Aug 2011	15 Apr 2012	244
Denmark	28 Aug 2011	17 Apr 2012	233
England (UK)	1 Aug 2011	31 Mar 2012	243
Estonia	1 Aug 2011	30 Apr 2012	273
Finland	30 Aug 2011	5 Apr 2012	219
Flanders (Belgium)	19 Aug 2011	31 Mar 2012	225
France	7 Sep 2012	24 Nov 2012	79
Germany	1 Aug 2011	31 Mar 2012	243
Ireland	1 Aug 2011	31 Mar 2012	243
Italy	1 Sept 2011	15 Apr 2012	227
Japan	30 July 2011	29 Feb 2012	214
Korea***	26 Sept 2011	24 Apr 2012	132
Netherlands	22 Aug 2011	11 May 2012	263
Northern Ireland (UK)	1 Aug 2011	13 Apr 2012	256
Norway	17 Aug 2011	30 Apr 2012	257
Poland*	1 Aug 2011	31 Mar 2012	243
Russian Federation* [8]	21 Nov 2011	29 May 2012	190
Slovakia	27 Oct 2011	24 Apr 2012	180
Spain	2 Sept 2011	30 Apr 2012	241
Sweden	22 Aug 2011	1 June 2012	284
United States	25 Aug 2011	3 Apr 2012	222

Source: Data Collection Form submitted after data cleaning unless otherwise noted.
* Based on Data Collection Form submitted after conclusion of data collection.
** Canadian PIAAC data collection was scheduled so as to not conflict with Census field activities.
*** Data collection was suspended due to administrative consideration between 23 December 2011 and 12 March 2012.

[7] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[8] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

10.3 Production and response rates

This section presents data on each country's production in terms of completes and response rates.

Table 10.2: Target and Actual Number of Completed Cases and Response Rates

	Number of Completed Cases		Overall Response Rate (without Reading Components)	
	Target	**Actual**	**Target**	**Actual**
Australia	7,928[3]	7,428	80	71
Austria	5,000	5,130	51	53
Canada	--[4]	27,285[3]	65	59
Cyprus[9]	4,500	5,053	67	73
Czech Republic	6,000[3]	6,102	70	66
Denmark	6,800[3]	7,328	60	50
England (UK)	5,000	5,131	56	59
Estonia	7,500[3]	7,632	61	63
Finland	5,150	5,464	64	66
Flanders (Belgium)	5,000	5,463	48	62
France	5,460	6,993	52	67
Germany	4,925	5,465	51	55
Ireland	5,600	5,983	64	72
Italy	4,455	4,621	51	56
Japan	5,000	5,278	50	50
Korea	5,000	6,667	80	75
Netherlands	5,000	5,170	50	51
Northern Ireland (UK)	3,600	3,761	56[2]	65
Norway	5,000	5,128	59	62
Poland	9,041[1,3]	9,366	57[1]	56
Russian Federation[10]	5,000[1]	3,892	54[1]	52
Slovakia	5,568[3]	5,723	65	66
Spain	5,876	6,055	45	48
Sweden	5,000	4,469	50	45
United States	5,000	5,010	68	70

Source: Data Collection Form submitted after data cleaning and Survey Design International File, unless otherwise noted.

[1] Based on Data Collection Form submitted after conclusion of data collection.
[2] A specific response rate target for Northern Ireland (UK) was not reported so it was assumed to be the same as for England.
[3] Country with oversamples and/or special populations.
[4] Not reported.

[9] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[10] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

(See Chapter 14 to understand why the targets for completes are different across the countries.) As shown in Table 10.2, all countries except Sweden met the PIAAC target number of completes required by the TSG. Sweden only completed 4,500 of the 5,000 completed cases normally required when administering problem solving.

The TSG requires countries to achieve a 70% overall response rate. However, the TSG also indicates that a response rate of 50% or better is acceptable if the results of a nonresponse bias analysis (when necessary) determine no significant bias is present. As noted above, at the end of this round of data collection, 22 countries had achieved a response rate of 50% or better.

In addition, for planning purposes, countries were asked to estimate their "expected" response rates at the beginning of the study. Seventeen countries met or exceeded their estimated target (Austria, Cyprus,[11] England-UK, Northern Ireland-UK Estonia, Finland, Flanders (Belgium), France, Germany, Ireland, Italy, Japan, Netherlands, Norway, Slovakia, Spain, and the United States).

To provide some context, Table 10.3 presents a comparison of PIAAC response rates to prior adult literacy surveys response rates – IALS and ALL – and to past rounds of the European Social Survey (ESS). This is provided for information purposes only as response rates are calculated differently across studies. PIAAC response rates are not directly comparable to past literacy studies such as IALS and ALL because of stringent restrictions on the level of exclusions permitted in PIAAC.

More details about sample sizes and response rates can be found in Chapter 16.

[11] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 10.3: PIAAC Overall Response Rates Compared to Other Recent International Surveys

	European Social Survey					IALS[6]		ALL[7]	PIAAC
	2002[1]	2004[2]	2006[3]	2008[4]	2010[5]	1994	1998	2003	2012
Australia	--	--	--	--	--	96.0	--	--	71
Austria	60.4	62.4	64.0	--	--	--	--	--	53
Canada	--	--	--	--	--	69.0	--	66.0	59[3]
Cyprus[12]	--	--	67.3	78.7	69.7	--	--	--	73
Czech Republic	43.3	55.3	--	69.5	70.2	--	61.5	--	66
Denmark	67.6	64.2	50.8	53.9	55.4	--	65.7	--	50
England (UK)	55.5	50.6	54.6	55.8	56.3	63.0	--	--	59
N. Ireland (UK)	55.5	50.6	54.6	55.8	56.3	63.0	--	--	65
Estonia	--	79.1	65.0	57.4	56.2	--	--	--	63
Finland	73.2	70.7	64.4	68.4	59.5	--	69.1	--	66
Flanders (Belgium)	59.2	61.2	61.0	58.9	53.4	36.0	--	--	62
France	43.1	43.6	46.0	49.4	47.1	--	--	--	67
Germany	55.7	51.0	54.5	48.0	30.5	69.0	--	--	55
Ireland	64.5	62.5	56.8	51.6	58.3	60.0	--	--	72
Italy	43.7	59.3	--	--	--	--	--	44.0	56
Japan	--	--	--	--	--	--	--	--	50
Korea[2]	--	--	--	--	--	--	--	--	75
Netherlands	67.9	64.3	59.8	49.8	60.0	45.0	--	--	51
Norway	65.0	66.2	65.5	60.4	58.0	--	60.9	56.0	62
Poland	73.2	73.7	70.2	71.2	70.3	75.0	--	--	56
Russian Federation[13]	--	--	69.5	67.9	66.6	--	--	--	52
Slovakia	--	62.7	73.2	72.6	74.7	--	--	--	66
Spain	53.2	54.9	65.9	66.8	68.5	--	--	--	48
Sweden	69.5	65.4	65.9	62.2	51.0	60.0	--	--	45
United States	--	--	--	--	--	60.0	--	66.0	70

Source: Data Collection Form submitted after data cleaning and Survey Design International File, unless otherwise noted.

[1] ESS1 - 2002 Summary and deviations. http://ess.nsd.uib.no/ess/round1/deviations.html.
[2] ESS2 - 2004 Summary and deviations. http://ess.nsd.uib.no/ess/round2/deviations.html.
[3] ESS3 - 2006 Summary and deviations. http://ess.nsd.uib.no/ess/round3/deviations.html.
[4] ESS4 - 2008 Summary and deviations. http://ess.nsd.uib.no/ess/round4/deviations.html.
[5] ESS5 - 2010 Summary and deviations, http://ess.nsd.uib.no/ess/round5/deviations.html.
[6] "Literacy in the Information Age." Final report of the International Adult Literacy Survey. Table B.6a and B.6b, p. 119,
http://www.oecd.org/education/highereducationandadultlearning/41529765.pdf.
[7] "Learning a Living." First results of the Adult Literacy and Lifeskills Survey, Table B8, p. 327,
http://www.oecd.org/education/educationeconomyandsociety/34867438.pdf .

[12] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[13] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

10.4 Interviewer training

Providing timely, adequate, standardized training to interviewers is an important tool in promoting the collection of quality data. Interviewers need to be very familiar with the survey procedures to administer them consistently across respondents and to produce data as error-free as possible. Familiarity with survey procedures allows interviewers to focus on gaining respondent cooperation, which in turn should help maximize response rates.

Chapter 9 of the TSG is dedicated entirely to training. It covers a variety of aspects associated with a successful training approach. Countries were, at a minimum, expected to:

- Conduct interviewer training in person, very close to the start of data collection.

- Train interviewers in small groups of 15-20.

- Assemble training staff to include a knowledgeable lead trainer, assistant(s), and technical support.

- Offer an adequate level of training. Although the Consortium recommended that countries should offer the same amount of training to all interviewers regardless of their level of experience, guidance was provided to tailor training to the level of experience of interviewers. (About 15 hours was recommended for experienced interviewers and 30 hours for interviewers new to survey research.)

- Provide sufficient hours of in-person training on **BQ and Direct Assessment administration** in the form of scripted mock interviews in which interviewers take turns reading the questions and a respondent (trainer or other interviewer) provides scripted answers. (About four hours recommended for each.)

- Provide sufficient hours of in-person training on **gaining cooperation** in the form of lectures and roundtable exercises where experienced interviewers are placed in groups with less experienced interviewers to discuss effective strategies for dealing with reluctant respondents. (About four hours recommended.)

10.4.1 Training logistics

The Consortium's recommendation was to conduct interviewer training the week before the start of data collection so interviewers could quickly apply the techniques learned and minimize learning loss. As shown in Table 10.4, most countries (68%) conducted interviewer training one to two weeks prior to the beginning of data collection. A significant proportion of countries (32%) held interviewer training sessions three weeks or more prior to data collection (Canada, France, Ireland, Italy, Japan, Netherlands, Norway and Poland). These countries typically organized several training sessions staggered in time so that only a fraction of interviewers received their training just before beginning work; for the first groups of interviewers to be trained, there was a considerable lag between training and data collection.

Six countries (24%) continued to train interviewers long after the start of data collection (more than four months) by organizing supplemental training sessions to compensate either for interviewer attrition or insufficient initial staffing.

A total of 380 interviewer training sessions were held worldwide, with numbers of sessions per country ranging from two in the Russian Federation[14] to 72 in Canada. The duration of training sessions varied significantly within and across countries. For example, the Netherlands held training that lasted between one and two days, while sessions held by Ireland lasted six to seven days.

Table 10.4: Summary of Main Study Interviewer Training Logistics

	Interviewer Training				
	Date Began	**Date Ended**	**Data Collection Start Date**	**Number of Sessions Held**	**Number of Days Per Event**
Australia	28 Sept 2011	25 Jan 2012	1 Oct 2011	15	3
Austria	11 July 2011	11 Nov 2011	1 Aug 2011	8	2-3
Canada	3 Oct 2011	6 Apr 2012	1 Nov 2011	72	4-5
Cyprus[15]	23 Aug 2011	7 Dec 2011	1 Sept 2011	9	2
Czech Republic	12 Aug 2011	14 Jan 2012	15 Aug 2011	15	2-3
Denmark	25 Aug 2011	11 Sept 2011	28 Aug 2011	4	2-4
England (UK)	18 July 2011	18 Nov 2011	1 Aug 2011	26	2
Estonia	12 July 2011	15 Dec 2011	1 Aug 2011	11	2-4
Finland	16 Aug 2011	7 Sept 2011	30 Aug 2011	7	2
Flanders (Belgium)	16 Aug 2011	18 Nov 2011	19 Aug 2011	7	3
France	4 July 2012	5 Sept 2012	7 Sept 2012	63	3
Germany	18 July 2011	12 Aug 2011	1 Aug 2011	5	3-5
Ireland	23 June 2011	28 July 2011	1 Aug 2011	3	6-7
Italy	22 June 2011	29 Sept 2011	1 Sept 2011	10	2-3
Japan	4 July 2011	29 July 2011	30 July 2011	14	4
Korea	15 Sept 2011	9 Mar 2012	26 Sept 2011	13	5
Netherlands	27 June 2011	12 Aug 2011	22 Aug 2011	16	1-2
Northern Ireland (UK)	25 July 2011	4 Nov 2011	1 Aug 2011	14	2
Norway	20 June 2011	30 Sept 2011	17 Aug 2011	12	2-5
Poland	6 July 2011	8 Feb 2012	1 Aug 2011	7	3
Russian Federation[16]	7 Nov 2011	2 Dec 2011	21 Nov 2011	2	3-4
Slovakia	6 Oct 2011	31 Jan 2012	27 Oct 2011	8	2
Spain	29 Aug 2011	2 Feb 2012	2 Sept 2011	29	3-4
Sweden	16 Aug 2011	2 Sept 2011	22 Aug 2011	6	1-3
United States	18 Aug 2011	13 Jan 2012	25 Aug 2011	4	4-6

Source: Interviewer Training Forms

[14] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[15] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[16] Please refer to the above note regarding the Russian Federation.

At each training session, countries were required to have at least one lead trainer, one assistant trainer and one person responsible for technical support. The lead trainer requirement was met by all countries. However, 52% of countries conducted some training sessions without an assistant and/or technical support staff (Table 10.5).

In addition, 17 countries (68%) exceeded the maximum number of 20 trainees per training room in some sessions.

Table 10.5: Interviewer Training Staffing and Class Sizes for the Main Study

	Number of Training Staff Per Session/Room			
	Lead	**Assist**	**Tech**	**Number of Trainees Per Session/Room***
Australia	1-2	0-4	1	2-22
Austria	2	1-4	2-3	9-26
Canada	1-2	0-1	0	1-26
Cyprus[17]	1-2	2-3	1-2	8-39
Czech Republic	3-4	0-2	1	8-21
Denmark	3-6	3-6	1	35-66
England (UK)	1-3	0-4	0-2	8-17
Estonia	5	3	5-6	7-20
Finland	2	5	1	11-23
Flanders (Belgium)	2-3	1	1	5-19
France	1-2	0-1	0-3	3-10
Germany	1-3	2-3	1	18-31
Ireland	2	1-2	2	15-23
Italy	2-3	0-4	0-3	14-22
Japan	1	0-2	1-2	9-23
Korea	2	1-2	0	2-58
Netherlands	4	1	4	Not reported
Northern Ireland (UK)	1	0-1	1	9-15
Norway	2-3	1-6	1-2	8-29
Poland	1	0-2	2-4	12-74
Russian Federation[18]	2	4	1	83-87
Slovakia	1-3	0-3	0-1	2-38
Spain	1-2	0-2	0-1	1-9
Sweden	5	1	1	20-24
United States	1-2	0-1	1-2	15-17

Source: Interviewer Training Form.

*A range indicates that a country conducted multiple training sessions with varying numbers of training staff and trainees. Only the minimum and maximum are reported here.

[17] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[18] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

10.4.2 Content covered

As shown in Table 10.6, the Consortium proposed an interviewer training program of approximately 15 hours for returning Field Test interviewers with good reviews (Profile 1) and 36 hours for new interviewers without any interviewing experience (Profile 4).[19] For interviewers with some experience on other surveys (Profile 3) or those with less than favorable Field Test reviews (Profile 2), the requirements were to essentially train interviewers as if they were new except that they could be exempted from training on administrative procedures and the case management system.

As countries were allowed to tailor their training program to their interviewers' particular needs, it is somewhat challenging to evaluate the adequacy of training offered. However, there were a certain number of topics for which virtually no tailoring was allowed for interviewers without PIAAC Field Test experience. These are BQ and assessment administration and gaining cooperation. For these topics (as well as others), the Consortium had provided detailed training materials that countries were required to use. As can be seen in Table 10.6, the time requirements in hours for these topics were essentially the same for Profiles 2, 3 and 4, that is, seven to 10 hours on BQ administration and 5.5 hours for assessment administration.

After completion of the Main Study, the Consortium realized that the BQ training materials required only four to five hours to be administered. Consequently, for the purpose of this report, countries were evaluated against this revised requirement. The duration of the assessment administration training was revised slightly to four hours.

[19] For countries using a screener, an additional two hours of training on screener administration was recommended.

Table 10.6: Required and Optional Components of Interviewer Training by Interviewer Profile (Main Study)

Interviewer Training Topic	Profile 1 PIAAC Field Test interviewers with good reviews	Profile 2 Interviewers with less than favorable PIAAC	Profile 3 Interviewers with some experience on other surveys	Profile 4 Interviewers without any interviewing experience
Introduction	0.50	0.50	1.75	1.75
Preparing for the field	0.50	0.50	0.50	0.50
CAPI training			1.50	1.50
Locating households/respondents		1.50	1.50	1.50
Case management system		(1.50)	(1.50)	1.50
Screener interactive, **if applicable**	1.25	1.25	2.00	2.00
BQ	4.00	7.00	8.00	10.00
Disposition codes	1.00	1.00	1.00	1.00
Direct Assessment	4.75	5.50	5.50	5.50
Core scoring	0.50	0.50	0.50	0.50
QC/Admin	(0.50)	(2.50)	(2.50)	2.50
Gaining respondent cooperation		4.50	2.50	4.50
Practice interview (role play)	1.75	1.75	3.25	3.25
Live respondent practice	2.00	2.00	2.00	2.00
Total hours for countries with list samples	**15.50 (15)**	**28.75**	**32 (28)**	**36**

Source: Clarifications Regarding Main Study Interviewer Training Requirements, 30 March 2011.

As shown in Table 10.7, fewer than half of countries (40%) met or exceeded the number of hours dedicated to gaining cooperation training (about four hours for new interviewers and two hours for those with prior experience; no gaining cooperation training was necessary for returning PIAAC Field Test interviewers with good reviews). Another 16% of countries met the requirement for some of the interviewers.

Only about half of countries (15, or 60%) spent the minimum recommended amount of time on BQ administration (four hours, regardless of level of experience). Another 17% met the requirement for some of their interviewers only. Only 14 (56%) of countries met the minimum number of hours required for assessment administration (about four hours). Another 17% of countries met the requirement partially (i.e., for some of their interviewers only).

Table 10.7: Actual Training Time Spent on BQ Administration, Assessment Administration and Gaining Cooperation, by Training Type and Interviewer Profile (Main Study)

	Training Type*	Number of Trainees	Majority Profile in Group	Hours In Person	BQ Total**	DA Total***	Coop Total
Australia	N/A	220	varied	19.75	4.5	3.6	.7
Austria	Full	50	3	28	5.5	6	1.5
	Reduced	97	varied	16.5	2	3.5	.5
Canada	N/A	759	varied	37.5	9	8.5	7
Cyprus[20]	N/A	150	4	18	2.5	2.8	1.5
Czech Republic	Full	159	varied	16.1	2.7	2.2	3
	Reduced	49	3	12	2.2	2.5	1.5
Denmark	Full	155	3	26	5	4	3.5
	Reduced	56	1	15	2.5	2	1
England (UK)	N/A	328	varied	10	1.3	3.5	.8
Estonia	Full	70	varied	33	7	4	4
	Reduced	43	3	24	5	4	1
	Reduced	19	3	17	4	3	0
Finland	N/A	124	1	15	4	2	2
Flanders (Belgium)	N/A	101	3	24	5	4	3
France	N/A	508	varied	18	4	3	.3
Germany	Full	91	3	31	6	7.8	3
	Reduced	38	1	22.3	3.5	6.3	3
Ireland	Full	38	3	44.5	6	8	5.5
	Reduced	18	3	38	4	7	4.5
Italy	N/A	170	varied	27	7	5	0
Japan	N/A	205	varied	23.8	2.8	3.3	1
Korea	N/A	229	3	30	7	6	3
Netherlands	Full	165	3	14.5	3.5	2.5	1
	Reduced	100	1	7.5	1	1.5	.5
Northern Ireland (UK)	N/A	165	3	10	1.3	3.5	.8
Norway	Full	42	4	19	2	5	0
	Reduced	15	1	14	0	6	4
	Reduced	98	1	16	1	3	1
Poland	N/A	236	3	25	6	4.5	2
Russian Federation[21]	Full	87	3	34.6	7.1	6.1	2.2
	Reduced	83	3	31.2	6.3	5.6	1.5
Slovakia	N/A	105	varied	20.4	5.5	3.5	.9
Spain	N/A	113	varied	18	3.0	4.0	1.7

[20] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[21] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 10.7 (cont.): Actual Training Time Spent on BQ Administration, Assessment Administration and Gaining Cooperation, by Training Type and Interviewer Profile (Main Study)

	Training Type*	Number of Trainees	Majority Profile in Group	Hours In Person	BQ Total**	DA Total***	Coop Total
Sweden	Full	64	3	17.9	4	3.8	1.5
	Reduced	68	1	7.6	0	2.3	1.5
United States	Full	186	3	38.3	4.8	5.3	5.5
	Full	3	3	32	4.3	4.5	4
	Full	6	3	31.8	4.5	5	2.5

Source: Interviewer Training Form.

* Several groups of interviewers may have been offered the same training type. For conciseness, groups offered the same training are not listed separately in this table.
** Includes time spent at in-person training on Introduction to CI/BQ administration, BQ interactives, and BQ exercises.
*** Includes time spent at in-person training on Introduction to Direct Assessment, Direct Assessment interactives, and Core Scoring.

The data in Table 10.7 suggest that several countries made significant adaptations to interviewer training scripts provided by the Consortium. Countries were permitted to make adaptations to Consortium training materials to fit their specific situation (mostly BQ adaptations), but these adaptations were not expected to dramatically affect the time spent on training.

The recommended amount of time to spend on BQ and assessment administration was deemed necessary for interviewers to get exposure to each question and become comfortable with the instruments. Interviewers must be unhindered by the technical aspects of survey administration to be able to focus on one of the most challenging part of their job – obtaining and maintaining cooperation from respondents. Spending significantly less time than recommended on these critical topics may have negatively affected response rate and/or data quality in many countries.

10.5 Staffing and field management

Hiring a sufficient number of fieldworkers (supervisors and interviewers), close supervision, and monitoring of production goals and response rates are fundamentals of successful fieldwork.

10.5.1 Interviewer hiring and attrition

Each country was required to hire a sufficient number of interviewers to achieve that country's production goals in eight months (see Table 10.2 for production targets). Because the optimal number of interviewers depends on numerous country-specific factors, the Consortium could not determine the exact number each country needed. However, TSG 8.3.1 provided specific considerations for countries. National teams were advised to use the best information available from similar national surveys conducted in their country as well as interviewers' PIAAC Field Test experience. Countries with compressed data collection schedules were advised to adjust their staffing needs accordingly.

Table 10.8 provides detailed information about staffing levels and attrition and suggests that countries learned from their Field Test experience and adjusted their staffing for the Main Study.

Twenty-three countries hired more than 100 interviewers (between 102 to 786 interviewers). Only two geographically small countries hired fewer than 100 – Ireland (61) and Cyprus[22] (84).

About 40% of countries experienced substantial levels of interviewer attrition (above 20%). All but four countries (88%) had some interviewer resignations. About 10 countries (40%) laid off interviewers, and 64% dismissed interviewers due to poor productivity or quality control issues.

Table 10.8: Data Collection Staffing and Attrition

	Number of Interviewers					Causes of Attrition				
	Attended Training	Received Assignment	Working at the End of Study	Typical Hours Worked Per Week	Attrition Rate (%)	Quit	Laid off	Productivity	Quality Control	Other
Australia	229	229	189	15-30	17	x				x
Austria	151	150	142	15	5	x	x	x		
Canada	810	786	274	5-25	65	x	x	x		
Cyprus[23]	150	84	5	20-40	94	x	x	x	x	
Czech Republic	194	194	74	20-40	62	x	x	x	x	
Denmark	216	216	192	8-20	11	x	x	x		x
England (UK)	343	328	243	10-25	26	x	x			
Estonia	127	124	75	30-40	40	x		x		
Finland	124	124	122	15-20	2	x				
Flanders (Belgium)	102	102	35	20	66	x		x		x
France	508	508	506	2-20	≈0					x
Germany	129	129	125	--**	3		x			
Ireland	70	61	40	25	34	x		x		
Italy	170	170	159	25-35	6	x		x	x	
Japan	228	226	224	5-35	1			x		x
Korea	220	220	216	40	2	x				
Netherlands (The)	275	275	167	10-15	39	x	x	x		
Northern Ireland (UK)	186	186	181	10	3	x				
Norway	140	140	134	10-25	4	x				
Poland*	286	286	196	18	31	x				x
Russian Federation* [24]	170	140	140	15-42	0	x			x	
Slovakia	107	107	97	8	9	x	x	x		x
Spain	144	139	117	30-40	16	x		x	x	
Sweden	145	137	135	10-15	2				x	x
United States	195	192	50	25-40	74	x	x	x	x	

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

* Based on Data Collection Form submitted after conclusion of data collection.
** Not reported.

[22] The number of interviewers hired by countries depended on several factors. For example, most countries had interviewers working part time while others had interviewers working full time on PIAAC (see Table 10.8 for the typical number of hours worked by PIAAC interviewers in each country).
[23] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[24] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

10.5.2 Field management

Two key indicators of adequate field management are: (1) the supervisor-interviewer ratio and (2) the frequency/regularity of supervisor-interviewer meetings.

In terms of the **interviewer-supervisor ratio,** countries were advised to assign one supervisor for every 15-20 interviewers to support the close supervision and mentoring of data collection. Table 10.9 indicates that 16 countries (64%) adhered to the recommended ratio of 20:1. However, when the ratio is increased to 30:1, only one country (Netherlands) stands out as far exceeding the Consortium recommendation with a ratio of 55:1.

Table 10.9: Number of Interviewers per Supervisor

	Number of Interviewers Who Received Assignments	Number of Supervisors	Size of Supervisor Assignment
Australia	229	10	15-22
Austria	150	6	27
Canada	786	80	8-10
Cyprus[25]	84	4	10-20
Czech Republic	194	6	15-25
Denmark	216	8	20-30
England (UK)	328	63	1-20
Estonia	124	8	11-15
Finland	124	6	20-30
Flanders (Belgium)	102	4	25
France	508	44	6-20
Germany	129	8	15-25
Ireland	61	4	12-14
Italy	170	10	10-20
Japan	226	31	2-20
Korea	220	61	2-5
Netherlands (The)	275	5	55
Northern Ireland (UK)	186	20	10
Norway	140	7	15-20
Poland*	286	50	2-6
Russian Federation* [26]	140	24	5-20
Slovakia	107	6	12-16
Spain	139	18	4-12
Sweden	137	6	23
United States	192	11	16-19

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

*Based on Data Collection Form submitted after conclusion of data collection.

[25] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[26] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

With regard to adequate *communication between field staff*, the TSG calls for weekly phone or in-person communication among the various levels of field staff and email communication as necessary. In particular, field supervisors should have weekly calls with their interviewers to ensure steady and adequate progress in data collection by keeping all staff on task, and making them accountable for their progress or lack thereof. Discussion during the meetings should focus on progress through caseload, response rates, problems encountered, and strategies/solutions for the completion of their remaining cases. Meeting sporadically can result in failures to meet data quality and production goals.

The majority of countries (16, or 64%) followed communication recommendations. Another six countries either had meetings every other week or less often (Finland, Poland, the Russian Federation[27]) or had variation across regions (Canada, Slovakia, and Spain). Only three countries had no scheduled meetings and opted to have meetings only as needed (Austria, Czech Republic, the Netherlands).

Countries used a variety of modes to communicate with their field staff. All countries used phone and all countries, but two used email (Denmark and Slovakia). Other strategies such as in-person meetings and newsletters were used by slightly more than half of countries. Some countries mentioned the use of newer technologies such as an online forum and video conferencing.

Details regarding the modes and frequency of communication are presented in Table 10.10.

Table 10.10: Modes of Communication Used between Field Staff during Data Collection

	Modes of Communication Used					
Country	**In Person**	**Phone**	**Email**	**Newsletter**	**Other**	**Frequency**
Australia		x	x	x	Lotus Notes database	Weekly
Austria	x	x	x			As needed
Canada	x	x	x	x		Varies
Cyprus[28]	x	x	x		Secure FTP Server, web service	Daily
Czech Republic		x	x	x		As needed
Denmark		x				As needed, weekly
England (UK)	x	x	x	x		As needed, weekly
Estonia	x	x	x		Online forum	As needed, weekly
Finland	x	x	x	x	Online forum	As needed, biweekly
Flanders (Belgium)	x	x	x			As needed, weekly
France	x	x	x			As needed, weekly
Germany		x	x	x		As needed, weekly
Ireland	x	x	x	x	Group briefing every 2 months	As needed, weekly

[27] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[28] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 10.10 (cont.): Modes of Communication Used between Field Staff during Data Collection

	Modes of Communication Used					
Country	**In Person**	**Phone**	**Email**	**Newsletter**	**Other**	**Frequency**
Italy	x	x	x		Video conferencing	As needed, weekly
Japan	x	x	x		Fax, message of Main Study	As needed, weekly
Korea	x	x	x		Q&A on the website	As needed, 2-3 times a week
Netherlands (The)		x	x	x		As needed, daily if necessary
Northern Ireland (UK)		x	x	x		As needed, weekly
Norway	x	x	x	x		As needed, weekly
Poland*	x	x	x			As needed, biweekly
Russian Federation* [29]		x	x	x	Video conferencing	Biweekly
Slovakia		x		x		Varies
Spain	x	x	x	x	Agency website	Varies
Sweden		x	x	x		Weekly
United States	x	x	x			As needed, weekly

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

* Based on Data Collection Form submitted after conclusion of data collection.

10.6 Nonresponse reduction

Implementation of a comprehensive strategy to promote survey participation is a key element in obtaining acceptable and/or high response rates. Doing so requires the use of a variety of outreach materials and tools, the collection of information on contact attempts and nonresponse, and strategies specifically aimed at minimizing nonresponse. In addition, countries were strongly advised, but not required, to offer a respondent incentive as a way to increase participation.

10.6.1 Use of introductory materials and outreach tools

Countries were required to send an introductory letter to households/respondents in advance of the interviewer visit and were advised to use a variety of tools to increase the visibility and legitimacy of the study. Table 10.11 shows that virtually all countries used an introductory letter, a study brochure, a study-specific website, and a respondent help line.[30] Endorsement letters, newspaper articles and press releases were used by about half of countries. Few countries made use of radio or TV advertisements.

[29] Please refer to the above note regarding the Russian Federation.

[30] This is a telephone line that potential respondents can call to receive additional information about the survey. The number for this line is usually provided in the introductory letter or the study brochure.

With regard to the use of respondent help lines by potential respondents, Table 10.11 shows that countries received widely varying numbers of calls. Among countries providing counts, Estonia received the fewest, with 20 calls, and Korea received the most, with 1,739 calls.

In addition, some countries participated in TV shows, held press conferences, and placed ads on the web and social media.

Table 10.11: Introductory Materials Used in Presenting the Study to Respondents/Households

	Intro. Letter	Study Brochure	Endorsement Letter	Newspaper Article	TV Ads	Radio Ads	Press Release	Study-Specific Website	Respondent Helpline (# calls)	Other
Australia	**x**	**x**						**x**	**x (n.r.)**	
Austria	**x**	**x**						**x**	**x (400)**	
Canada	**x**	**x**						**x**	**x (1491)**	
Cyprus[31]	**x**	**x**	**x**	**x**			**x**	**x**	**x (133)**	
Czech Republic	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x (386)**	
Denmark	**x**	**x**		**x**		**x**	**x**	**x**	**x (505)**	
England (UK)	**x**	**x**						**x**	**x (823)**	
Estonia	**x**	**x**		**x**	**x**	**x**	**x**	**x**	**x (20)**	posters, video, web ads
Finland	**x**	**x**	**x**	**x**			**x**	**x**		TV show, social network ads
Flanders (Belgium)	**x**	**x**						**x**	**x (375)**	
France	**x**	**x**							**x (500)**	letters to mayor's office and police stations
Germany	**x**	**x**	**x**	**x**			**x**	**x**	**x (307)**	flyers
Ireland	**x**	**x**	**x**	**x**			**x**	**x**	**x (115)**	
Italy		**x**	**x**	**x**			**x**	**x**	**x (168)**	press conference
Japan	**x**	**x**	**x**	**x**	**x**		**x**	**x**	**x (1644)**	
Korea	**x**	**x**	**x**					**x**	**x (1739)**	posters, banners
Netherlands (The)	**x**	**x**						**x**	**x (400)**	
Northern Ireland (UK)	**x**	**x**						**x**	**x (242)**	
Norway	**x**	**x**		**x**			**x**	**x**	**x (912)**	Main Study messages
Poland*	**x**	**x**		**x**			**x**	**x**	**x (90)**	refrigerator magnet
Russian Federation* [32]	**x**	**x**	**x**					**x**		
Slovakia	**x**	**x**		**x**		**x**	**x**	**x**	**x (90)**	call-back cards
Spain	**x**	**x**						**x**	**x (198)**	letters to local councils/condos
Sweden	**x**	**x**		**x**			**x**	**x**	**x (n.r.)**	radio/TV interviews
United States	**x**	**x**	**x**					**x**	**x (183)**	refrigerator magnet, tailored flyers, pens

Source: Data Collection Form submitted after data cleaning unless otherwise noted.
* Based on Data Collection Form submitted after conclusion of data collection.

31 Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

32 Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

10.6.2 Documenting contact attempts

Countries were advised to require interviewers to thoroughly document each contact attempt with sample persons/households and to record as much information as possible on nonresponse cases. The purpose was to provide supervisors with the information necessary to manage the work of interviewers effectively and help them be productive.

The information recorded for each contact attempt had to include, at a minimum, the date, time and outcome of each visit. Interviewers were also supposed to provide comments that might prove helpful in obtaining respondent cooperation during future contacts.

Table 10.12 provides a summary of the information recorded by countries about each contact attempt and nonresponse cases. All countries recorded all elements recommended about contact attempts. However, a significant number of countries (n=7) did not provide an opportunity for interviewers to write comments about the case, which can be very helpful when planning nonresponse work.

Table 10.12: Information Collected by Interviewers about Contact Attempts during Main Study Data Collection

	Day	Date	Time	Mode	Outcome	Comments	Other
Australia	x	x	x	x	x		
Austria	x	x	x	x	x		
Belgium	x	x	x	x	x	x	
Canada	x	x	x	x	x	x	Several other
Cyprus[33]	x	x	x	x	x	x	
Czech Republic	x	x	x	x	x	x	
Denmark	x	x	x	x	x		
England (UK)	x	x	x	x	x	x	
Estonia	x	x	x	x	x	x	
Finland	x	x	x	x	x	x	
France	x	x	x	x	x	x	
Germany	x	x	x	x	x	x	
Ireland	x	x	x	x	x	x	
Italy	x	x	x	x	x	x	
Japan	x	x	x	x	x		
Korea	x	x	x	x	x	x	
Netherlands	x	x	x	x	x		
Northern Ireland (UK)	x	x	x	x	x	x	
Norway	x	x	x	x	x	x	Interviewer ID
Poland*	x	x	x	x	x	x	
Russian Federation* [34]	x	x	x	x	x	x	
Slovakia	x	x	x	x	x		
Spain	x	x	x		x	x	
Sweden	x	x	x	x	x		
United States	x	x	x	x	x	x	Interviewer ID

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

* Based on Data Collection Form submitted after conclusion of data collection.

[33] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[34] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

10.6.3 Monitoring contact attempts

At a minimum, countries were required to ensure that a minimum number of contact attempts were made to each respondent/household. Countries were strongly advised to attempt initial contacts in person and to make at least three subsequent contact attempts. Countries for which telephone initial contacts are customary were allowed to deviate from this standard but were required to make at least six subsequent attempts.

Table 10.13 presents details of the contact procedures used by participating countries. It shows that 21 countries (84%) used in-person initial contacts, either exclusively or in combination with the telephone. An additional three countries (France, Italy and the Russian Federation[35]) used a hybrid strategy in which some individuals were initially contacted by personal visits and others by telephone. All countries met the minimum number of contacts required with respect to their mode choice.

Table 10.13: Strategy for Contacting Potential Respondents/Households during Main Study Data Collection

	Mode of Initial Contact		Minimum Number of Subsequent Contacts	
	In Person	**Telephone**	**In Person**	**Telephone**
Australia	**x**		5	5
Austria	**x**		4	0
Canada	**x**		5	20
Cyprus[36]	**x**		4	5
Czech Republic	**x**		5	0
Denmark	**x**		5	0
England (UK)	**x**		6	0
Estonia	**x**		7	2
Finland		**x**	4	0
Flanders (Belgium)	**x**		5	0
France	**x**	x	5	7
Germany	**x**		4	0
Ireland	**x**		4	0
Italy	**x**	**x**	4	7
Japan	**x**		4	0
Korea	**x**		4	7
Netherlands	**x**		6	0
Northern Ireland (UK)	**x**		3	0
Norway		**x**	3	7
Poland*	**x**		4	0
Russian Federation* [37]	**x**	**x**	4	7
Slovakia	**x**		4	0
Spain	**x**		6	4
Sweden		**x**	0	10
United States	**x**		4	0

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

[35] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[36] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[37] Please refer to the above note regarding the Russian Federation.

[*] Based on Data Collection Form submitted after conclusion of data collection.

Finland, Norway and Sweden used the telephone as the sole mode for initial contact, although in-person visits were used by Finland and Norway to supplement the telephone strategy. Sweden made no in-person visits, except when interviewers could do so on the way to an appointment.

10.6.4 Documenting nonresponse

In addition to recording information about each contact attempt, countries were also required to record details about each case finalized as nonresponse. These included basic demographics about the person who refused, the strength of the refusal, the likelihood of conversion, any problems encountered, and any relevant information that might facilitate future contact with a potential respondent.

The level of detail recorded varied from country to country. However, all countries recorded basic information about nonrespondents, as shown in Table 10.14.

Table 10.14: Information Collected by Interviewers on Nonresponse Cases during Main Study Data Collection

	Demographics	Refusal Strength	Problems Encountered	Conversion Likelihood	Comments	Other
Australia	x[2]	x	x	x		
Austria	x	x	x	x	x	
Canada		x	x	x	x	x[3]
Cyprus[38]	x	x	x	x	x	
Czech Republic	x	x	x	x	x	
Denmark	x	x	x			
England (UK)	x	x	x	x	x	x[6]
Estonia	x	x	x	x	x	
Finland	x	x		x	x	x[4]
Flanders (Belgium)	x	x	x	x	x	
France	x	x	x	x		
Germany		x	x		x	x[5]
Ireland		x	x	x	x	
Italy	x	x	x		x	
Japan	x	x	x	x	x	
Korea	x	x	x	x	x	
Netherlands	x	x	x	x	x	
Northern Ireland (UK)	x	x	x	x	x	
Norway	x				x	
Poland[1]	x		x		x	
Russian Federation[1] [39]		x	x	x	x	
Slovakia	x	x	x	x	x	
Spain	x	x	x	x		
Sweden	x	x	x	x	x	
United States	x	x	x	x	x	x[7]

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

[1] Based on Data Collection Form submitted after conclusion of data collection.
[2] Only if screener has been completed.
[3] Refusals are escalated to supervisor and manager level for resolution, and these steps are recorded in the case management system.
[4] Type of refusal.
[5] Presence of an intercom, house type, condition of the house, respondent's social class and education as appraised by the interviewer prior to first contact attempt.
[6] In one of the data collection agencies: recommendation for profile of interviewer who is more likely to be successful at converting the case.
[7] Name and phone number of a contact person.

In addition, countries were asked to report some of the most common reasons for refusal to do the BQ (Table 10.15) and assessment (Table 10.16). For the BQ, lack of interest was the most

[38] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[39] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

often cited reason across all countries, followed by lack of time (“too busy”). For the assessment, the excessive length (“too long”) and lack of time were the most often cited refusal reasons.

Table 10.15: Most Common Reasons for Refusal to BQ (Main Study)

	Not Interested	Too Long	Don’t Want To Be Bothered	Don’t Trust Surveys	Too Busy	Other
Australia		x			x	
Austria	x					
Canada		x	x		x	
Cyprus[40]	x	x			x	
Czech Republic	x		x		x	
Denmark	x	x	x			
England (UK)	x		x		x	
Estonia	x		x		x	
Finland	x		x		x	
Flanders (Belgium)	x	x			x	
France	x	x	x			
Germany						x[2]
Ireland	x		x		x	
Italy	x	x		x	x	x[3]
Japan	x				x	x[4]
Korea	x	x	x			
Netherlands	x	x	x			
Northern Ireland (UK)	x		x		x	
Norway	x	x			x	
Poland[1*]	x		x		x	
Russian Federation[1*] [41]	x				x	
Slovakia	x		x		x	
Spain	x		x		x	
Sweden	x				x	x[5]
United States	x	x			x	

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

[1*] Based on Data Collection Form submitted after conclusion of data collection.
[2] Legal guardian refused respondents’ participation; respondent doesn't want to give more information (altogether three refusals).
[3] Literacy-related problems.
[4] Sickness, poor physical condition.
[5] Voluntary nature of the survey.

[40] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[41] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 10.16: Most Common Reasons for Refusal to Assessment (Main Study)

	Not Interest-ed	Too Long	Don't Want To Be Bothered	Waste of Time and Money	Too Busy	Don't Want To Do Exercise	Too Complicated	Other
Australia		x			x			
Austria		x				x	x	
Canada		x			x	x		
Cyprus[42]		x			x			
Czech Republic						x		
Denmark		x				x	x	
England (UK)	x	x			x			
Estonia		x			x	x		
Finland		x				x	x	
Flanders (Belgium)		x			x	x		
France		x			x	x		
Germany			x			x	x	
Ireland		x			x		x	
Italy		x			x	x	x	
Japan								
Korea		x				x	x	
Netherlands		x					x	
Northern Ireland (UK)					x			
Norway								
Poland*		x		x	x			
Russian Federation* [43]	x	x			x			
Slovakia	x		x		x			
Spain	x		x		x			
Sweden	x				x			x**
United States	x	x			x			

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

* Based on Data Collection Form submitted after conclusion of data collection.
** Voluntary nature of the survey.

10.6.5 Nonresponse strategy

Countries were strongly advised to implement a comprehensive strategy to deal with nonresponse cases. Most countries (92%) implemented a strategy involving a combination of techniques, such as case reassignment, senior interviewer follow-up and the use of tailored letters. Two countries had strategies involving only the use of case reassignment (UK-Northern Ireland) or supervisor follow-up combined with tailored letters (Korea). However, Korea and Northern Ireland (UK) offered substantial monetary incentives (64 and 37 Euros, respectively), and secured response rates at or above 65%. Table 10.17 presents each strategy in detail.

[42] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[43] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 10.17: Strategies to Deal with Difficult/Nonresponse Cases during Main Study Data Collection

	Case Re-Assign.	Follow-Up Senior FIs	Follow-Up Supervisors	Traveling Reassignment	Tailored Letters	Refusal Conversion Letters	Other
Australia	x	x			x	x	
Austria	x	x				x	x[2]
Canada	x	x	x		x	x	
Cyprus[44]	x		x	x	x	x	
Czech Republic	x	x		x	x	x	
Denmark	x	x	x			x	
England (UK)	x	x	x	x		x	
Estonia	x				x	x	
Finland	x	x			x	x	
Flanders (Belgium)	x	x	x			x	
France			x		x	x	
Germany	x	x			x	x	x[3]
Ireland	x		x	x	x	x	
Italy	x		x	x			x[4]
Japan	x	x	x		x	x	x[5]
Korea			x		x	x	x[6]
Netherlands	x	x		x	x	x	
Northern Ireland (UK)	x						
Norway	x	x		x	x	x	
Poland[1*]	x	x	x			x	
Russian Federation[1*] [45]	x		x		x		x[7]
Slovakia	x	x				x	x[8]
Spain	x		x		x	x	
Sweden	x	x	x	x	x	x	x[9]
United States	x	x	x	x	x	x	x[10]

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

1* Based on Data Collection Form submitted after conclusion of data collection.
2 Motivation calls to nonrespondents of certain groups identified in order to reduce bias.
3 In certain cases, interviewers had access to funding for discretionary incentives (only symbolic) or received day rates (e.g., for refusal conversion).
4 Refusal conversion calls.
5 Offered option to conduct interview at home or out of home such as community hall. Designed mobile and PC websites to allow respondents to schedule appointment for interview.
6 Field managers or field directors tried to persuade some respondents.
7 Contact leaders of local communities and ethnic diasporas; contact building managers.
8 Telephone calls to the households by field managers, supervisors.
9 Group of interviewers dedicated to refusal conversion.
10 Tailored flyers, mail-in screener forms sent to sampled households yet to be screened.

44 Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

45 Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

10.6.6 Use of incentives

The vast majority of countries (23, or 92%) offered some form of incentive. Two countries (8%), Australia and Canada, have rules preventing the use of incentives in government surveys. Among countries that offered an incentive, 18 (78%) used a monetary incentive. Details regarding the nature of each country's incentive are provided in Table 10.18.

Table 10.18: Respondent Incentives Used During Main Study Data Collection

	Incentive type		
	Monetary**	**Non-Monetary**	**Description**
Australia			None
Austria	x		50 EUR shopping voucher
Canada			None
Cyprus[46]	x		50 EUR shopping voucher from popular furniture store
Czech Republic	x		500 CZK (approx. 20 EUR)
Denmark		x	100 DKR (approx. 13 EUR) to respondents who participated in live practice. Lottery ticket in the last phase of the data collection period
England (UK)	x	x	30 GBP (approx. 37 EUR) voucher; booklet of stamps as a refusal conversion tool in some areas.
Estonia		x	Magazine subscription
Finland		x	USB flash drive; lottery of popular tablet computer
Flanders (Belgium)		x	Lottery ticket (3 EUR)
France		x	Numeracy kit
Germany	x	x	Study-specific adhesive notepad sent to all with introductory letter and brochure; 50 EUR upon completion
Ireland	x		30 EUR shopping voucher
Italy	x		30 EUR shopping coupon (increased to 40 EUR in the last 10 weeks of fieldwork).
Japan		x	Book voucher
Korea	x		4 EUR for completed screener + 20 EUR for completed BQ + 40 EUR for completed assessment
Netherlands	x		20 EUR voucher (increased to 40 EUR in the final stage of data collection)
N. Ireland (UK)	x		30 GBP (approx. 37 EUR) voucher
Norway	x	x	Refrigerator magnet to all; 500 NOK (approx. 66 EUR) gift card upon completion
Poland*	x	x	8 EUR shopping voucher; lottery ticket
Russian Federation* [47]	x		300 RUB (approx. 7 EUR) or 500 RUB (approx. 12 EUR) depending on regions
Slovakia	x		10 EUR
Spain	x	x	Choice of 20 EUR voucher or equivalent donation to NGO
Sweden	x	x	Refrigerator magnet to all; 10 EUR check upon completion
United States	x	x	Study-specific refrigerator magnet and pen to all; 50 USD upon completion (approx. 40 EUR)

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

* Based on Data Collection Form submitted after conclusion of data collection.

[46] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[47] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

** The distinction between monetary and nonmonetary incentive is somewhat subjective. Strictly speaking, anything other than cash or a check is not monetary. However, "shopping vouchers" were considered to be monetary incentives unless they could only be exchanged for very specific goods such as books or magazine subscriptions.

10.7 Fieldwork quality control

Each country was required to implement a comprehensive fieldwork quality control plan. This plan had to include:

- an evaluation of interviewing skills at the end of training and during data collection
- the implementation of a validation (callback) effort to detect falsification
- the review of survey and process data through the analysis of automated management reports

10.7.1 Audio recording/observation of interviews

Countries were strongly advised to monitor at least two interviews per interviewer during the early stages of data collection and provide feedback. Monitoring could either be done by audio recording interviews, observing the interviews in person, or a combination of both.

The vast majority of countries (22, or 88%) did some form of monitoring. Among these 22 countries, 15 monitored at least one interview per interviewer on average, but few reached the recommended level of two interviews per interviewer (see ratio of interviews monitored to number of interviewers assigned in Table 10.19). The Consortium's recommendation was to monitor the second and 10th complete achieved by each interviewer. However, some interviewers may not have been productive enough to allow for a country to monitor a second interview. Therefore, countries are considered to have met the standard if they have monitored at least one interview per interviewer on average. Ten countries did not meet this reduced standard. Australia, Austria and Northern Ireland (UK) did not monitor any interviews. Canada, Finland, France, Italy, the Netherlands, Norway and England (UK) did not monitor enough interviews given the number of interviewers they assigned to PIAAC.

Table 10.19 shows the number of interviewers assigned to PIAAC, the number of interviews that were audio recorded or observed in each country, and the ratio of interviews monitored to the number of interviewers assigned to PIAAC work.

Table 10.19: Number of Interviews Monitored by Mode during the Main Study Data Collection

	Number of Interviewers Assigned	Number of Interviews Monitored				Ratio of Interviews Monitored to Number of Interviewers Assigned
		Taping Full Interview	Taping Snippets	Observation	Total	
Australia	229	0	0	0	0	0
Austria	150	0	0	0	0	0
Canada	786	0	0	385	385	0.49
Cyprus[48]	84	121	0	0	121	1.44
Czech Republic	194	0	199	0	199	1.03
Denmark	216	440	0	0	440	2.04
England (UK)	328	0	0	41	41	0.13
Estonia	124	503	0	0	503	4.05
Finland	124	101	0	0	101	0.81
Flanders (Belgium)	102	135	0	0	135	1.32
France	508	0	0	400	400	0.79
Germany	129	245	4	0	249	1.93
Ireland	61	100	0	40	140	2.29
Italy	170	0	0	165	165	0.97
Japan	226	0	425	0	425	1.88
Korea	220	682	0	218	900	4.09
Netherlands	275	36	0	0	36	0.13
Northern Ireland (UK)	186	0	0	0	0	0
Norway	140	0	0	120	120	0.86
Poland*	286	1800	0	0	1800	6.29
Russian Federation* [49]	140	1250	0	0	1250	8.93
Slovakia	107	0	306	0	306	2.86
Spain	139	176	44	0	220	1.58
Sweden	137	274	0	0	274	2.00
United States	192	298	0	0	298	1.55

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

* Based on Data Collection Form submitted after conclusion of data collection.

10.7.2 Validation

Validation (back-checks) is critical to data validity; it is the most important QC feature of household data collection. As stated in the TSG, the validation procedure required the verification of "10% of an interviewer's finalized work, including cases finalized as nonresponse." The validation plan had to ensure that:

- validation cases were selected randomly;
- at least 10 percent of each interviewer's cases were validated; and
- all dispositions were validated, not just completes

[48] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[49] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 10.20 presents a summary of the validation procedures implemented by countries.

Table 10.20: Summary of Adherence to Validation Standards by Countries (Main Study)

Requirements	Implementation
Cases must be selected randomly, not ad hoc	20 countries (80%) selected most or all cases randomly
At least 10% of each interviewer's work must be validated	11 countries (44%) reached the 10% threshold for each interviewer 17 countries (68%) reached a threshold of 7% for 90% of interviewer
All types of disposition must be validated (completes, non-contacts, refusals, ineligibles)	22 countries (88%) validated all dispositions

Source: Data Collection Form submitted after data cleaning. For Poland and the Russian Federation,[50] the source is the Data Collection Form submitted after conclusion of data collection.

The requirement to validate *each* interviewer at the 10% level appears to have been the most challenging for countries to meet, as only 11 countries did so. Even when setting the threshold lower (7% of cases validated for 90% of interviewers), only 17 countries met this requirement.

Regarding other validation requirements, 20 countries selected most or all validation cases randomly (Germany, Japan, Poland and England (UK) only selected *some* cases randomly; France didn't select any cases randomly) and 22 countries (88%) validated all dispositions (Australia and Japan did not validate cases finalized as ineligible; France only validated cases finalized as completes).[51]

Details about each country's validation procedure are presented in Table 10.21.

[50] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[51] Based on information provided on QC forms and during monthly QC conference calls, the Russian Federation followed validation requirements. However, analysis of the data revealed evidence of falsification affecting a significant proportion of cases. This level of falsification should have been detected by validation. The fact that it was not suggests that validation was not conducted in a manner sufficiently adequate to uncover falsification.

Table 10.21: Summary of Validation Procedure for Main Study

	Percentage of interviewers…		Validation Mode				Dispositions Validated					
	validated at 10%	validated at 7%	Phone	In Person	Mail	Other	Completes	Non-Contacts	Refusal	Ineligible	Random Selection	Consortium Form Used
Australia	0	0	793	0	0	0	x	x	x		All	No
Austria	90	94	1122	0	0	0	x	x	x	x	All	Adapted
Canada	65	85	5357	0	0	0	x	x	x	x	All	Adapted
Cyprus[52]	100	100	637	2	379	0	x	x	x	x	All	As is
Czech Republic	100	100	2189	877	0	0	x	x	x	x	Most	Adapted
Denmark	100	100	990	0	7	0	x	x	x	x	Most	As is
England (UK)	12	20	524	2872	33	0	x	x	x	x	Some	Adapted
Estonia	98	100	1138	588	620	0	x	x	x	x	All	As is
Finland	16	46	559	0	0	0	x	x	x	x	All	Adapted
Flanders (Belgium)	75	84	1006	0	0	0	x	x	x	x	All	Adapted
France	100	100	0	0	6,684	0	x				No	No
Germany	100	100	175	176	3400	39	x	x	x	x	Some	Adapted
Ireland	100	100	918	275	12	0	x	x	x	x	Most	As is
Italy	96	99	1450	0	0	0	x	x	x	x	Most	Adapted
Japan	100	100	996	171	589	0	x	x	x		Some	Adapted
Korea	100	100	745	134	0	0	x	x	x	x	All	As is
Netherlands (The)	76	86	584	0	665	0	x	x	x	x	Most	Adapted
Northern Ireland (UK)	91	95	219	1124	2133	0	x	x	x	x	Most	Adapted
Norway	100	100	830	0	0	0	x	x	x	x	All	Adapted
Poland*	36	40	0	1499	0	0	x	x	x	x	Some	Adapted
Russian Federation* [53]	100	100	2500	0	0	0	x	x	x	x	All	As is
Slovakia	97	97	1708	140	0	0	x	x	x	x	Most	As is
Spain	100	100	1045	320	0	0	x	x	x	x	Most	Adapted
Sweden	80	91	860	0	230	0	x	x	x	x	All	Adapted
United States	100	100	1611	228	54	0	x	x	x	x	Most	As is

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

* Based on Data Collection Form submitted after conclusion of data collection.

[52] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[53] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Thirteen countries (52%) uncovered instances of falsification involving one to 22 interviewers. The nature of falsifications was as follows:

- underreporting the number of household members in the screener
- completing the BQ over the telephone
- using a proxy respondent for the BQ
- misusing the disposition codes
- leaving the exercise booklets with the respondent overnight
- having someone other than the selected respondent complete the exercise
- making up answers to the BQ and the exercise

This emphasizes the critical importance of validation for in-person studies in which interviewers are working independently in the field. A rigorous validation procedure is critical to substantiating data quality.

10.7.3 Other quality control checks

Countries were advised to use automated management reports (proposed by the Consortium) dealing with process data as well as any other means of detecting falsification available to them. The majority of countries (88%) used some of the reports proposed by the Consortium to monitor administration length, time lapse between interviews, and the number of interviews completed per day. Three countries (France, Japan and Russian Federation[54]) did not. Details are provided in Table 10.22.

Table 10.22: Use of Fieldwork Quality Control Reports During the Main Study Data Collection

	Interview Duration	Individual Instrument Duration	Time Between Interviews	Interviews Conducted Very Late/Very Early	Number of Interviews Per Day	Other
Australia	x	x				x[2]
Austria	x	x	x	x	x	
Canada	x	x	x	x	x	
Cyprus[55]	x	x	x	x	x	
Czech Republic	x	x	x	x	x	
Denmark	x	x	x	x	x	
England (UK)	x	x	x	x	x	
Estonia	x	x	x	x	x	
Finland	x		x	x	x	
Flanders (Belgium)	x	x	x	x	x	
France						
Germany	x	x	x	x	x	x[3]
Ireland	x	x	x	x	x	x[4]
Italy	x		x	x		
Japan						
Korea	x	x	x	x	x	
Netherlands	x			x	x	
Northern Ireland (UK)	x	x	x	x	x	
Norway	x	x	x	x	x	x[5]
Poland[1*]	x	x	x	x	x	x[6]
Russian Federation[1*] [56]						
Slovakia	x				x	
Spain	x	x	x	x	x	
Sweden	x	x			x	
United States	x	x	x	x	x	

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

[1*] Based on Data Collection Form submitted after conclusion of data collection.
[2] Number of calls and spread of days and times.
[3] Consistency checks of interview and register data (age, gender, nationality).
[4] Review of interviews conducted over 2 days.
[5] Population register checks.
[6] Inconsistency between some BQ items; respondent's actual and declared birthdate.

[54] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[55] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[56] Please refer to the above note regarding the Russian Federation.

In addition, countries were advised to monitor the quality of data throughout the Main Study data collection. Most countries (88%) reviewed data frequencies and missing data rates. All countries but France reviewed the quality of open-ended responses. Details are presented in Table 10.23.

Table 10.23: Procedures to Monitor Quality of Data During the Main Study Data Collection

	Data Frequencies	Review of 'Other-Specify' Responses	Review of Open-Ended Responses	Missing Data Rates	Other
Australia	x	x	x		
Austria	x		x	x	x[2]
Canada	x	x	x	x	x[3]
Cyprus[57]	x	x	x	x	
Czech Republic	x	x	x	x	
Denmark	x	x	x	x	
England (UK)		x	x	x	
Estonia	x	x	x	x	
Finland	x	x	x	x	
Flanders (Belgium)	x	x	x	x	
France	x			x	
Germany	x		x	x	
Ireland	x	x	x	x	
Italy			x		
Japan		x	x	x	
Korea	x	x	x	x	
Netherlands		x	x	x	
Northern Ireland (UK)	x	x	x	x	
Norway	x	x	x	x	
Poland[1*]	x	x	x	x	
Russian Federation[1*] [58]	x	x	x	x	
Slovakia	x		x		
Spain	x	x	x	x	
Sweden	x	x	x	x	x[4]
United States	x	x	x	x	

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

[1] Based on Data Collection Form submitted after conclusion of data collection.
[2] Consistency checks.
[3] Cross-tabulations, merging of files for consistency checks, fixing data discrepancies.
[4] Macro checks of data; distributions of select background variables have been checked against distribution of corresponding variable from population register and Labor Force Survey.

[57] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[58] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

10.7.4 Interviewer productivity

Unusually high interviewer productivity (measured in number of completed interviews) can be an indication of falsification. Countries were asked to monitor the minimum, mean and maximum number of completes per interviewer and to increase the validation rate of interviews with high production.

The mean number of completes per interviewer ranged from 15 in England (UK) to 102 in Ireland. These countries were characterized by an unusually large and unusually small interviewer workforce, respectively, which is reflected in these numbers. In most countries, the mean number of completes per interviewer was in the 30-40 range. The maximum ranges were extremely varied from 51 in the Russian Federation[59] to 317 in Spain. Details about interviewer productivity are presented in Table 10.24.[60]

Table 10.24: Summary Statistics of the Number of Completes Achieved by Interviewers for the Main Study Data Collection

	Mean (rounded to nearest unit)	Minimum	Maximum	Range
Australia	34	1	123	122
Austria	35	10	116	106
Canada	33	1	132	131
Cyprus[61]	47	0	199	199
Czech Republic	34	1	177	176
Denmark	36	1	130	129
England (UK)	15	1	52	51
Estonia	60	3	195	192
Finland	45	14	91	77
Flanders (Belgium)	56	1	272	271
France	--**	--**	--**	--**
Germany	41	8	82	74
Ireland	102	11	156	145
Italy	26	1	97	96
Japan	23	3	73	70
Korea	34	15	58	43
N. Ireland (UK)	20	1	64	63
Netherlands	20	1	137	136
Norway	45	3	143	140
Poland*	39	1	138	137
Russian Fed.* [62]	35	5	51	46
Slovakia	56	1	159	158

[59] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[60] Interviewer productivity may have been influenced by the number of hours worked (see Table 10.8 for the typical number of hours worked by PIAAC interviewers in each country).

[61] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[62] Please refer to the above note regarding the Russian Federation.

Table 10.24 (cont.): Summary Statistics of the Number of Completes Achieved by Interviewers for the Main Study Data Collection

	Mean (rounded to nearest unit)	Minimum	Maximum	Range
Spain	42	1	317	316
Sweden	35	4	89	85
United States	25	1	123	122

Source: Data Collection Form submitted after data cleaning unless otherwise noted.

[1] Based on Data Collection Form submitted after conclusion of data collection.
[2] Not reported.

10.8 Interviewer debriefing

Countries were required to administer a debriefing questionnaire and/or hold debriefing sessions with interviewers to gain insights into the problems they encountered, comments made by respondents, and suggestions for improving procedures for future cycles of PIAAC. Countries were required to provide a report to the Consortium summarizing the key findings. An international summary covering interviewer training, BQ and Exercise administration, and the virtual machine was provided to all Consortium members and OECD.

10.9 Recommendations for future cycles

The Field Test and Main Study provided opportunities for countries, the Consortium and the OECD to understand country compliance with the implementation of PIAAC according to a set of agreed-upon standards and to meet production goals. Based on the Field Test and Main Study experience of PIAAC Round 1, the Consortium is proposing a series of recommendations for future cycles of PIAAC.

1. **Study planning is crucial to success, and timely submission of the National Survey Design and Planning Report (NSDPR) must be a nonnegotiable requirement for participation.** Completing a thorough and timely NSDPR ensures: (1) that countries have thought through the study requirements/challenges and are prepared to assure the Consortium that they are fully committed to PIAAC; and (2) allows the Consortium to provide timely feedback on areas of concern.

   In particular, countries hiring data collection organizations separate from the survey institute must be confident that their contractor intends to meet the TSG and can provide all the information necessary to submit a complete and timely NSDPR.

2. **Countries must conduct a rigorous survey institute selection and monitoring process.** Countries should start the search and selection process for the organization that will undertake data collection as early as possible. Final selection should occur no later than six months prior to the start of data collection.

Countries should provide candidate organizations with clear experience and performance guidelines based on the TSG. Final selection should be based on demonstrated experience and the ability to perform the work following the PIAAC TSG.

Countries must monitor the data collection entity closely during the period of performance, requiring at least monthly meetings with key organization staff as well as monthly reports. During the data collection period, countries should also require weekly production status reports.

3. **All cycles of PIAAC must include a Field Test.** Cycle 1 countries learned a great deal from the Field Test experience, which allowed them to adjust their data collection process in preparation for the Main Study. Due to expected changes in future cycles, especially in terms of content and country staffing, all future cycles of PIAAC should require a Field Test, even for countries having implemented a successful Cycle 1.

4. **The Main Study data collection period should be extended one to two months.** This will accommodate countries dealing with populations that have fairly rigid holiday observances (e.g., August vacations, winter breaks, religious periods) and experience weather/climate challenges to plan accordingly and meet the data collection timeframe.

5. **Countries should adhere to the training program produced by the Consortium and train field staff following the TSG hours specified.** Successful data collection requires interviewers that are well trained on the importance of the survey, instrument administration and procedures, and obtaining high response rates. The guidelines below help ensure that country interviewers receive sufficient training. These guidelines for training hours have been revised based on the Cycle 1 experience (i.e., reduced by six hours) and are displayed in Table 10.25.

Table 10.25: Revised Interviewer Training Requirements in Hours

Topics	Inexperienced Interviewers	Experienced Interviewers
General Interviewing Techniques	4	--
Protocol		
- procedures	8	8
- screener administration	4[1]	4[1]
- BQ+EX administration	6	6
- role-plays/practice interviews	7	7
Gaining Cooperation	3	3
Total	28-32 hours	24-28 hours

[1] Not applicable for countries with registry samples.

6. **Countries should adopt a rigorous field management style as specified by the TSG.** Close management of fieldwork is crucial to the success of data collection. Countries must require continual monitoring of field staff and an adequate supervisor to interviewer ratio. NPMs should require country data collection managers to communicate weekly with them and their field staff to ensure adequate monitoring of production and response

rates. Data collection staff at all levels, from supervisors to interviewers, must be held accountable for their performance. This can be best achieved through frequent communication and monitoring.

7. **All validation TSG must be followed.** Countries must be required to agree to adhere to these standards, with no exceptions. This is the most important quality control activity undertaken in household studies. Thus, validation cases must be randomly selected from a sample of all finalized cases and must be conducted at the 10% level on all interviewers working on PIAAC.

Chapter 11: Quality Control Monitoring Activities

Pat Montalvan and Michael Lemay, Westat

11.1 Overview

This chapter presents the details of the survey operations' quality assurance (QA) and quality control (QC) activities conducted by the Consortium as part of the Main Study.

This program was designed to: (1) support the collection of data and results that are valid, reliable and comparable across countries and over time, and satisfy accepted quality assurance goals; (2) keep the OECD and the Consortium informed about the progress of data collection; and (3) provide advice to countries needing assistance. The aim was to implement a program that represents the best tradeoff of quality and cost within the constraints of the project. The plan was presented to the OECD and the BPC and approved by the PIAAC Technical Advisory Group (TAG) in July 2008.

The principal objectives of the QA and QC program for survey operations/data collection were the following:

- Undertake activities which monitor the implementation of the PIAAC Technical Standards and Guidelines (TSG) for survey operations during the Field Test and Main Study.
- Review the progress of data collection and identify potential problems and solutions on a timely basis during the Field Test and Main Study.
- Make recommendations to enhance the Main Study based on the Field Test experience.
- Identify sources of nonsampling error to inform analysis.
- Make recommendations for the next wave/cycle of PIAAC.

The PIAAC QC process provided continuous support to countries in following the TSG before, during and after data collection. It assisted countries by answering questions and communicating areas of concern in a timely manner. Furthermore, it informed OECD and the Consortium of the status of data collection in each country on a regular basis throughout the process. The level of cooperation from countries was superior to the Field Test and was very good overall.

The process described in this chapter allowed collection of most of the information presented in Chapter 10.

A description of the QA and QC activities for survey operations follows in section 11.2. Section 11.3 looks at country compliance with these activities.

11.2 Quality assurance and quality control activities

11.2.1 Quality assurance (QA)

The QA process for survey operations consisted of the development of standards and guidelines, including the QC process, a QC and management manual, and the training of national teams on QC activities.

Development of standards and guidelines for survey operations

The first step in the implementation of the PIAAC quality assurance program was the specification of standards and guidelines covering all aspects of the survey life cycle, from sample and instrumentation design to data processing and analysis. A significant portion of the TSG (Chapters 8 through 10) deals specifically with survey operations concerns such as field staff recruitment, management and training, and field management practices. The PIAAC standards are based on generally agreed upon policies or best practices to be adhered to in the conduct of the survey.

Development of survey operations QC and management manual

The purpose of this manual was to: (1) provide national teams with details on important survey operations standards with practical suggestions on how to implement them (e.g., field management reports, fieldwork quality control, tools to increase respondent cooperation); and (2) provide national teams with details on the logistics of the PIAAC quality control program (e.g., forms to be submitted, quality control call schedule).

International training on survey operations QC

The international training on survey operations QC control took place prior to the Field Test international interviewer training in February 2010 and covered the essential points in the QC manual. Key points were covered again at the June 2011 NPM meeting prior to the Main Study data collection.

11.2.2 Quality control (QC)

The QC process consisted in regular communication in the form of reports, conference calls and ad hoc email exchanges. This section provides a summary description of each activity.

National Survey Design and Planning Report (NSDPR) review

Each country was required to submit an NSDPR covering all aspects of the survey implementation at least six months prior to the beginning of data collection. The Consortium reviewed the survey operations chapters (four chapters, covering 70 standards) of the NSDPR for each country and reported on any serious deviations from the TSG.

Data collection QC conference calls

The Consortium conducted conference calls (see Chapter 5 in QC and Management Manual for more details) with each of the PIAAC countries on a regular basis throughout the critical Field Test and Main Study data collection periods. The goals of the calls were to: (1) review the content of the monthly data collection QC forms submitted by countries (see below); (2) give countries the opportunity to ask questions in real time; and (3) discuss any survey operations issues that may have arisen in each country.

Calls were held prior to the start of data collection, during data collection, and one month after data collection. Calls were held monthly with each country during Field Test data collection and reduced to every other month during the Main Study. However, calls were held more often when needed.

Conference call participants varied somewhat from month to month, depending on study timeframe and issues at hand, but generally they included the country's NPM, key Leading Survey Institute (LSI) staff (who speak English), and key Westat operations staff.

Conference calls followed a specific agenda guided by the data collection QC form and were documented using QC meeting minutes reports which summarized the items discussed, the decisions made and the pending action items.

Data collection QC form

Countries were required to complete monthly QC monitoring forms. These forms were used to guide the conference call meetings and focused on the topics covered in Chapters 2, 8, 9 and 10 of the PIAAC TSG (82 short answer questions). Topics included:

- field staffing and management
- plans for contacting households/respondents
- respondent outreach materials
- ways of dealing with nonresponse
- field management system
- response rates and production
- field QC measures
- plans to train staff on ethics and confidentiality

The completed electronic forms were posted each month on the project's SharePoint site, which is accessible by all participating countries and organizations.

Interviewer training QC form

To ascertain adherence to the interviewer training program designed by the Consortium, countries were required to complete an interviewer training QC form at the end of each interviewer training session in each country (28 questions). The form included questions about the:

- number of trainers and trainees
- experience profile of trainees
- training materials used
- topics covered at training

The completed electronic forms were also posted on the project's SharePoint site.

Interviewer debriefing questionnaire and report

Countries were required to administer a debriefing questionnaire to interviewers following the conclusion of data collection to ensure that interviewer feedback was obtained. The form included 47 questions covering:

- training
- the administration of the Background Questionnaire
- the administration of the computer-based exercise
- the administration of the paper exercise
- the interview in general
- the interviewer help line

Each country was required to summarize interviewer feedback for each question on the questionnaire and submit the report to the Consortium for review.

Ongoing Web communication

Through Web communication, countries could ask for and receive responses from Westat to ad hoc questions arising throughout the planning and implementation phases of PIAAC data collection.

11.3 Country compliance

As shown in Table 11.1, virtually all countries fulfilled the QC requirements for Main Study data collection. Some countries met the requirements with some delay but were proactive in notifying the Consortium in advance. A few calls had to be rescheduled, but this was usually done with advance notice.

Table 11.1: Compliance with the Main Study Survey Operations Quality Control (QC) Program

Required QC Activities	Percentage of Countries Complying (n=25)
Revised Main Study NDSPR (1 report)	96
QC calls	--
- at least once prior to data collection	96
- at least every other month during data collection	100
- once after data collection	100
Data Collection Form	--
- at least once prior to data collection	96
- monthly during data collection	84
- once after data collection	100
- once after data cleaning	88
Interviewer Training Form (1 form per training session)	100
Interviewer Debriefing Report (1 report)	88

Next, we report in detail how countries fulfilled the requirements.

Survey operations sections of the revised Main Study NSDPR

Twenty-four of the 25 participating countries (96%) submitted a final NSDPR for the Main Study, although few did so on time (by 1 February 2011). One country submitted only a draft Main Study NSDPR (see Table 11.2).

Table 11.2: Final Main Study NSDPR Submission Dates

Country	Submission Date
Australia	4 February 2011
Austria	15 March 2011
Canada	25 February 2011
Cyprus[1]	1 February 2011
Czech Republic	25 January 2011
Denmark	31 January 2011
England/N. Ireland (UK)	2 February 2011
Estonia	1 February 2011
Finland	31 January 2011
Flanders (Belgium)	26 January 2012
France	3 October 2012
Germany	1 February 2011
Ireland	9 August 2011
Italy	2 August 2011
Japan	31 January 2011 *(revised 9 March 2012)*
Korea	Draft Main Study only
Netherlands (The)	21 March 2012
Norway	1 February 2011 *(revised 12 August 2011)*
Poland	24 January 2011 *(revised 2 August 2011)*
Russian Federation[2]	1 February 2011
Slovakia	31 October 2011
Spain	11 February 2011 *(revised 27 April 2012)*
Sweden	10 February 2011
United States	1 February 2011

Source: PIAAC SharePoint site timestamps.

Data Collection Form submission and conference calls prior to data collection

Twenty-four countries (96%) submitted the required Data Collection Form and participated in a QC call at least once prior to the beginning of data collection, which is satisfactory. The requirement called for the submission of a Data Collection Form for each month leading up to the beginning of data collection. A few countries could not fulfill this requirement due to staff shortages during summer vacation. As in the Field Test, it appears that a few countries (n=9) may have misunderstood the requirement to submit a new form even if there were no changes.

[1] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[2] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

One country (Slovakia) did not submit a Data Collection Form and did not participate in a QC call prior to the beginning of data collection (see Table 11.3).

Table 11.3: Actual Schedule of Data Collection Form Submission and Associated QC Monitoring Calls Prior to Main Study Data Collection

	About Two Months Prior to Data Collection		About One Month Prior to Data Collection	
	Form	**Call**	**Form**	**Call**
Australia	5 August	16 August	12 Sept	20 Sept
Austria	7 June	7 June	Not submitted	Not required
Canada	8 Sept	20 Sept	Not required[2]	Not required
Cyprus[3]	18 July	26 July	16 August	Not required
Czech Republic	Not submitted	22 July	15 August	16 August
Denmark	1 July	6 July	27 July	3 August
England/Northern Ireland (UK)	26 May	6 June	29 June	Not required
Estonia	16 June	16 June	7 July	12 July
Finland	Not required[1]	Not required[1]	8 August	16 August
Flanders (Belgium)	23 June	27 June	23 August	Not required
France	Not required	Not required	13 July	27 July
Germany	Not required[1]	Not required[1]	4 July	13 July
Ireland	2 June	7 June	Not submitted	Not required
Italy	6 July	26 July	29 July	8 August
Japan	3 June	14 June	7 July	12 July
Korea	16 August	23 August	Not submitted	Not required
Netherlands (The)	20 July	21 July	Not submitted	24 August
Norway	20 June	22 June	Not submitted	Not required
Poland	22 June	22 June	25 July	25 July
Russian Federation[4]	5 August	10 August	Not submitted	Did not take place
Slovakia	Not submitted	Did not take place	Not submitted	Did not take place
Spain	8 July	15 July	5 August	23 August
Sweden	4 July	5 July	Not submitted	Not required
United States	20 July	Not required[3]	28 July	Not required[3]

Source: SharePoint and email timestamps, QC meeting minutes reports.

[1] A special agreement was reached in which it was agreed that the Data Collection Form submission and the QC call would take place in August only due to the difficulty of having staff available during summer vacations.
[2] It was agreed that a new submission was not necessary as the country certified that no change would be made to procedures.
[3] The Consortium's survey operations quality control manager attended weekly management meetings of the US PIAAC team.

Interviewer Training Forms

Twenty-five countries (100%) reported on their interviewer training sessions. The requirement was for countries to report on each training session held by submitting a separate report for each.

[3] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[4] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Worldwide, 380 interviewer training sessions were held. Countries conducted between two and 72 training sessions each, a number that includes both the initial training sessions and any session held to compensate for interviewer attrition.

Data Collection Form submission and conference calls during data collection

The majority of countries (84%) submitted one Data Collection Form for each month of fieldwork as required. All countries complied with the requirement to participate in a QC call at least every other month. The standard quality control program during the data collection period called for monthly submissions of the Data Collection Form (eight submissions) and QC conference calls at least every other month (at least four calls).[5] However, depending on their respective data collection start date and the end date of QC activities set by the Consortium (3 April 2012, extended to 2 May 2012), a number of countries had fewer than the typical number of submissions/calls. They are nonetheless considered to have fully complied with the quality control program (see Table 11.4).

[5] Requirements were adapted to France's shorter data collection period.

Table 11.4: Actual Schedule of Data Collection Form Submission and Associated QC Monitoring Calls for Main Study

	Month 1		Month 2		Month 3		Month 4		Month 5		Month 6		Month 7		Month 8	
	Form	**Call**	**Form**	**Call**	**Form**	**Call**	**Form**	**Call**	**Form**	**Call**	**Form**	**Call**	**Form**	**Call**	**Form**	**Call**
Australia	19 Oct	NR	8 Nov	14 Nov	13 Dec	20 Dec	11 Jan	NR	14 Feb	21 Feb	13 Mar	NR	--	--	--	--
Austria	30 Aug	8 Sept	20 Oct	NR	14 Nov	17 Nov	2 Dec	NR	10 Jan	12 Jan	16 Feb	NR	8 Mar	8 Mar	--	--
Canada	25 Nov	29 Nov	NS	NR	20 Jan	30 Jan	NS	NR	26 Mar	29 Mar	30 Apr	NR	--	--	--	--
Cyprus[6]	21 Sept	27 Sept	18 Oct	NR	14 Nov	22 Nov	13 Dec	NR	18 Jan	24 Jan	21 Feb	NR	20 Mar	27 Mar	--	--
Czech Republic	20 Sept	23 Sept	26 Oct	31 Oct	28 Nov	29 Nov	NR[3]	NR	25 Jan	27 Jan	21 Feb	NR	20 Mar	27 Mar	--	--
Denmark	7 Sept	NR	29 Sept	5 Oct	1 Nov	NR	30 Nov	7 Dec	10 Jan	NR	23 Jan	1 Feb	1 Mar	7 Mar	--	--
England/N. Ireland (UK)	31 Aug	7 Sept	28 Sept	NR	27 Oct	2 Nov	1 Dec	NR	21 Dec	4 Jan	25 Jan	NR	29 Feb	7 Mar	--	--
Estonia	4 Aug	NR	20 Sept	22 Sept	4 Nov	NR	4 Nov	8 Nov	6 Dec	13 Dec	9 Jan	NR	8 Feb	14 Feb	12 Mar	NR
Finland	14 Sept	NR	12 Oct	18 Oct	10 Nov	NR	13 Dec	20 Dec	11 Jan	NR	15 Feb	21 Feb	12 Mar	NR	--	--
Flanders (Belgium)	22 Sept	28 Sept	21 Oct	NR	23 Nov	28 Nov	15 Dec	NR	16 Jan	23 Jan	16 Feb	NR	14 Mar	2 Apr	--	--
France	NS	NR	3 Oct	4 Oct	NS	NR	--	--	--	--	--	--	--	--	--	--
Germany	2 Aug	NR	7 Sept	14 Sept	5 Oct	NR	2 Nov	9 Nov	7 Dec	NR	4 Jan	11 Jan	1 Feb	NR	8 Mar	14 Mar
Ireland	29 Aug	14 Sept	26 Oct	NR	7 Nov	9 Nov	13 Dec	NR	11 Jan	11 Jan	17 Feb	NR	14 Mar	14 Mar	--	--
Italy	10 Oct	17 Oct	14 Nov	NR	6 Dec	12 Dec	NS	NR	11 Feb	15 Feb	19 Mar	NR	--	--	--	--
Japan	14 Sept	14 Sept	5 Oct	NR	1 Nov	8 Nov	15 Dec	NR	6 Jan	10 Jan	7 Feb	NR	--	--	--	--
Korea	22 Oct	25 Oct	27 Nov	NR	17 Dec	21 Dec	NR[2]	NR[2]	NR[2]	NR[2]	30 Mar	3 Apr	--	--	--	--
Netherlands (The)	28 Sept	NR	14 Oct	20 Oct	10 Nov	NR	8 Dec	15 Dec	20 Jan	NR	10 Feb	16 Feb	12 Mar	NR	--	--
Norway	6 Sept	9 Sept	18 Oct	25 Oct	1 Dec	NR	21 Dec	21 Dec	25 Jan	NR	9 Feb	22 Feb	29 Mar	NR	--	--
Poland	8 Aug	NR	5 Sept	19 Sept	31 Oct	NR	21 Nov	21 Nov	8 Dec	NR	17 Jan	17 Jan	21 Feb	NR	20 Mar	20 Mar
Russian Federation[7]	28 Nov	22 Dec	23 Jan	25 Jan	25 Feb	28 Feb	28 Mar	NR	15 Apr	18 Apr	--	--	--	--	--	--
Slovakia	18 Nov	25 Nov	11 Dec	16 Dec	16 Jan	NR	21 Feb	23 Feb	26 Mar	NR	--	--	--	--	--	--
Spain	20 Sept	NR	14 Oct	21 Oct	14 Nov	NR	9 Dec	16 Dec	16 Jan	NR	10 Feb	17 Feb	9 Mar	NR	--	--
Sweden	28 Sept	4 Oct	NS	NR	15 Nov	6 Dec	17 Jan	NR	7 Feb	14 Feb	28 Mar	NR	--	--	--	--
United States	30 Aug	NR[1]	28 Sept	NR[1]	28 Oct	NR[1]	22 Nov	NR[1]	3 Jan	NR[1]	24 Jan	NR[1]	21 Feb	NR[1]	20 Mar	NR[1]

Source: SharePoint and e-mail timestamps and QC meeting minutes reports.

[1] The Consortium's survey operations quality control manager attended weekly management meetings of the US PIAAC team.
[2] Data collection was suspended.
[3] Not required by special agreement due to holiday break.

[6] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[7] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Data Collection Form submission and conference calls after data collection

Twenty-five countries (100%) submitted a Data Collection Form after completion of data collection. However, only 22 countries (88%) submitted a Data Collection Form after data cleaning was completed despite several reminders (see Table 11.5).

Table 11.5: Actual Schedule of Data Collection Form Submission and Associated QC Monitoring Calls

	After Data Collection		After Data Cleaning	
	Form	**Call**	**Form**	**Call**
Australia	13 Apr	17 Apr	21 June	Not required
Austria	30 Mar	12 Apr	11 June	Not required
Canada	27 July	Did not take place[1]	17 Sept	Not required
Cyprus[8]	17 Apr	24 Apr	25 May	Not required
Czech Republic	23 Apr	27 Apr	27 June	Not required
Denmark	2 Apr	4 Apr	6 July	Not required
England/N. Ireland (UK)	30 Mar	4 Apr	6 July	Not required
Estonia	4 Apr	10 Apr	19 June	Not required
Finland	11 Apr	17 Apr	20 June	Not required
Flanders (Belgium)	16 Apr	23 Apr	7 Aug	Not required
France	21 Dec	11 Jan	Not submitted	Not required
Germany	5 Apr	16 Apr	27 June	Not required
Ireland	17 Apr	18 Apr	31 Jan 2013	Not required
Italy	26 Apr	2 May	3 July	Not required
Japan	13 Mar	13 Mar	15 June	Not required
Korea	24 Apr	25 Apr	9 Aug	Not required
Netherlands (The)	22 Apr	26 Apr	18 June	Not required
Norway	25 Apr	25 Apr	20 June	Not required
Poland	12 Apr	16 Apr	Not submitted	Not required
Russian Federation[9]	6 June	Did not take place[1]	Not submitted	Not required
Slovakia	25 Apr	26 Apr	15 June	Not required
Spain	13 Apr	20 Apr	18 June	Not required
Sweden	28 Mar	3 Apr	23 Aug	Not required
United States	27 Apr	Not required[2]	15 June	Not required[2]

Source: SharePoint and e-mail timestamps, QC meeting minutes reports.

[1] Main Study quality control calls ended on 31 May 2012 for all countries but France.
[2] The Consortium's survey operations quality control manager attended weekly management meetings of the US PIAAC team.

[8] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[9] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Interviewer Debriefing Report

Countries were required to debrief their interviewers on their Main Study experience and provide a report to the Consortium. Twenty-two countries (88%) submitted such a report (see Table 11.6).

Table 11.6: Main Study Interviewer Debriefing Report Submission Dates

Country	Date
Australia	4 June
Austria	20 July
Canada	6 September
Cyprus[10]	15 May
Czech Republic	20 June
Denmark	3 August
England/N. Ireland (UK)	8 October
Estonia	19 June
Finland	29 May
Flanders (Belgium)	30 May
France	Not submitted
Germany	4 July
Ireland	20 June
Italy	18 June
Japan	10 September
Korea	9 August
Netherlands	30 July
Norway	28 August
Poland	Not submitted
Russian Federation[11]	17 July
Slovakia	7 June
Spain	27 June
Sweden	Not submitted
United States	18 May

Source: SharePoint and e-mail timestamps.

[10] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[11] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

11.4 Conclusion

Overall, the PIAAC quality control program for survey operations met its goals. During the Main Study: (1) country compliance was high; (2) the OECD and the Consortium were kept informed about the progress of data collection; (3) countries were supported – their questions were answered and areas of concern were pointed out promptly throughout the critical months before and during data collection; (4) the program allowed for the sharing of status information with all countries and helped foster a sense of cooperation and "shared experience"; and (5) the program experience should assist counties and the OECD as they plan for future cycles of PIAAC.

Chapter 12: Scoring Reliability Studies

Claudia Tamassia, Mary Louise Lennon and Kentaro Yamamoto, ETS

While PIAAC was primarily a computer-delivered and computer-scored instrument, a paper-and-pencil version of the cognitive instruments was also an important component of the assessment. The Field Test design required all participating countries to administer paper-based versions of the literacy and numeracy items as part of the study to compare the performance of items that served to link the paper-and-pencil and computer-delivery formats.[1] In the Main Study, paper booklets were administered to study participants who were either unwilling to take the test on the computer or unable to do so because they lacked basic computer skills. Therefore scoring designs and operational procedures were developed for these human-scored items.

This chapter describes the scoring process and associated scoring reliability studies for the paper-and-pencil instruments. Without accurate and consistent scoring of paper-and-pencil items, all subsequent psychometric analyses of those items are severely jeopardized. Therefore PIAAC, like other large-scale assessments before it, defined a set of essential processes that all participating countries were required to implement to maintain scoring consistency within and across countries. These included having items scored independently by two different scorers and providing a common set of anchor booklets to be scored by all national teams. An important aspect related to scoring in PIAAC was the requirement that countries follow specified scoring designs to ensure that each booklet was scored twice and that scorers functioned in both the first- and second-scorer roles across all the booklets. These scoring designs, along with a specified set of procedures for training scorers and monitoring the scoring process, were designed to ensure that PIAAC would provide accurate and reliable data for policymakers, researchers, and other stakeholder groups interested in adult skills and their distribution in an international context.

12.1 The scoring process

The PIAAC paper instruments included four booklets:

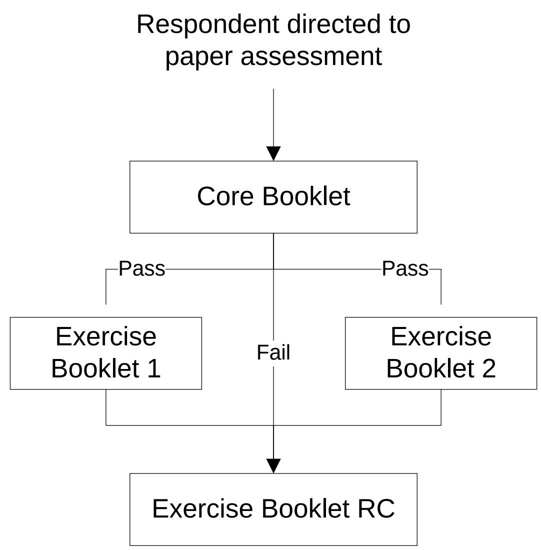
- the Core Booklet, which contained eight items (four literacy and four numeracy),
- Exercise Booklet 1, which contained 20 literacy items,
- Exercise Booklet 2, which contained 20 numeracy items, and

[1] Because the ICT component was an integral part of problem solving in technology-rich environments, there was no paper-based version of that domain.

- Exercise Booklet RC, which contained 109 reading components items.

In the Main Study, the paper-based assessment was administered to respondents who either reported they had no computer experience; failed the test of basic computer skills required to take the assessment; or refused to take the assessment on the computer. Within this design, the Core Booklet was presented first and included the easiest questions. If the respondent passed the Core Booklet, the interviewer administered either Exercise Booklet 1 or Exercise Booklet 2. Each respondent had a 50% chance of receiving one or the other booklet. In countries that opted to assess reading components, after the respondent completed Exercise Booklet 1 or 2, or in cases where a respondent failed the core, the interviewer administered Exercise Booklet RC. This Main Study design is illustrated in Figure 12.1.

Figure 12.1. Main Study paper booklet assessment design



The Core Booklet, Exercise Booklet 1 and Exercise Booklet 2 were scored by trained teams within each participating country. These same booklets were used to conduct within- and across-country reliability studies as described in section 12.2.

Responses for reading components (Exercise Booklet RC) were not scored. Instead, members of the scoring team recorded answers on response sheets that were then used for data entry and automated scoring. Therefore, the PIAAC scoring designs include only the Core Booklet and Exercise Booklets 1 and 2.

12.1.1 Preparing for scoring

A number of key activities were completed by the Consortium and National Centers prior to the assessment to prepare for scoring the paper-based instruments. The Consortium developed detailed scoring guides that included scoring rubrics as well as examples of correct and incorrect responses. For linking items, scoring information from previous assessments (IALS and ALL) was included in the scoring guides. For new items, scoring rubrics were defined for the Field Test, and information from Field Test scoring was then used to expand the scoring guides for the Main Study.

A two-day meeting with NPMs and chief scorers was conducted where scoring guides were presented and explained. Participants practiced scoring sample items, and the group discussed any ambiguous or problematic situations. By focusing on sample responses likely to provide the greatest challenge to scorers, meeting participants had the opportunity to ask questions and clarify the application of scoring rubrics. When the discussion revealed areas where rubrics could be improved, those changes were made and included in an updated version of the scoring guide documents provided after the meeting.

To support countries during the scoring process, the Consortium established a distribution list that allowed national teams to submit scoring questions and receive responses from the relevant domain experts. National teams were also able to review questions submitted by other countries along with the responses from the Consortium. A summary report of scoring issues was provided on a regular basis and all emails were archived on a discussion board on the PIAAC SharePoint site for reference by national scoring teams.

National Centers were responsible for assembling a team of scorers. The first task was to identify a lead scorer who would be part of the scoring team and additionally be responsible for the following tasks:

- Hiring and training scorers within the country
- Monitoring the scoring process. This included daily monitoring of the scores in the data entry software (Data Management Expert, or DME), reviewing scoring progress and outcomes, and taking action when scoring anomalies were identified. At the beginning of the scoring process, the lead scorer was required to manually inspect a portion of the scored booklets for scoring accuracy before scores were entered into the DME. This series of checks ensured that the initial booklets were scored according to the guidelines. When the lead scorer was comfortable and confident that all the scorers were consistently following the scoring guidelines, he or she then monitored outcomes through the DME software.
- Monitoring the inter-rater reliability and taking action when the scoring results were unacceptable and required further investigation
- Retraining or replacing scorers if necessary
- Subscribing to the PIAAC scoring distribution list, submitting any scoring questions for resolution by the PIAAC domain experts, and monitoring the weekly summary reports
- Reporting scoring results and providing status reports to the NPM and Consortium

The lead scorer was required to be proficient in English, as international training and interactions with the Consortium were in English only. It was also assumed that the lead scorer for the Field Test would retain that role for the Main Study. When this was not the case, it was the responsibility of the National Center to ensure that the new lead scorer received training equivalent to that provided at the international scoring training prior to the Field Test.

The guidelines for assembling the rest of the scoring team included the following requirements:

- All scorers were to have more than a high school qualification, with university graduates preferable.

- Scorers were to be trained based on a nationally developed training package that included an overview of the survey and training manuals based on the manuals and materials provided by the Consortium.

- The lead scorer and one other scorer were required to be bilingual, meaning they had to be proficient in English and the national language. Both scorers would serve as part of the scoring team and be responsible for scoring the anchor booklets. If countries followed a design that required only two scorers, both had to be bilingual.

- Scorers were expected to be organized in teams and to work on the same schedule and in the same location to facilitate discussion about scoring issues as they arose. Past experience showed that if scorers were able to discuss questions among themselves and with their lead scorer, many issues could be resolved in a way that resulted in more consistent scoring.

- Each scorer was assigned a unique scorer ID.

- Due to normal attrition rates and unforeseen absences, the Consortium strongly recommended that lead scorers train a backup for their scoring teams.

Additional information about the scoring staff was provided in standard 11.4 in the PIAAC Technical Standards and Guidelines.

12.2 Reliability studies

Comparability both within and across countries was an important design criterion in PIAAC. The equivalence of scoring was established by double scoring the Core Booklet, Exercise Booklet 1 and Exercise Booklet 2 by two independent scorers, as well as carefully monitoring and responding to the scoring results. These steps were required as quality-assurance measures to determine whether scorers were applying the scoring rubrics consistently. The purposes for rescoring were to: i) document the degree to which the same scores were given to items regardless of the scorer; and ii) identify items and scorers with low inter-rater agreement. To ensure that the first and second scores were truly independent, certain precautions were taken. For example, scores had to be assigned by two different scorers, and the second scorer was not allowed to see scores given by the first scorer.

12.2.1 Within-country scoring reliability study

The purpose of the within-country inter-rater scoring reliability study was to ensure scoring reliability within a country and identify scoring inconsistencies or problems early in the scoring process so they could be resolved as soon as possible. In general, inconsistencies or problems were due to scorer misunderstanding of general scoring guidelines and/or a rubric for a particular item.

The level of agreement between two scorers was represented by an inter-rater reliability index based on percent correct. In PIAAC, inter-rater reliability represented the extent to which any two scorers agreed on how a particular response should be scored, and thus how comparably the scoring rubric was being interpreted and applied. Inter-rater reliability varied from 0 (no reliability or 0% agreement) to 1 (maximum degree of reliability or 100% agreement). The goal for PIAAC was to reach a within-country inter-rater reliability of 0.95 (95% agreement) across all items, with at least 85% agreement for each item.

The IEA DME Tools Software was developed for calculating inter-rater reliability. As the name implies, these tools were used with data from the DME database. Once scores were entered into the database, the IEA DME Tools were used to produce output and reports needed for examining scoring reliability. Countries received training on the use of these tools to monitor scoring reliability.

12.2.2 Cross-country scoring reliability study

Accurate and consistent scoring within a country does not necessarily imply that all countries are applying the scoring guides in the same manner. Scoring bias may be introduced if one country scores a certain response differently from other countries. Therefore, in addition to within-country inter-rater reliability, it was also important to check the consistency of scorers across countries.

Guidelines 11.3.3A and 11.3.3B in the PIAAC Technical Standards and Guidelines described the cross-country reliability study using a set of anchor booklets. The anchor booklets were a set of 180 completed core, literacy and numeracy booklets. Item responses in these booklets were based on actual responses collected in the field as well as responses that reflected key points on which scorers were trained. They included the international cover page and were identified by an international booklet serial number (or ID) prefilled on that cover page. The anchor booklets were in English and scored by the two bilingual scorers.

The anchor booklets were required to be scored and rescored by every country as the main and reliability scores for these booklets were used by the Consortium to calculate inter-rater agreement across countries. There was no scoring resolution for these booklets. Thus, countries were to simply single score these booklets and enter the data into the DME. It was important that countries did not resolve any discrepancies with the anchor booklet items because the Consortium needed the original scores to examine the degree of consistency among the participating countries.

12.3 Scoring designs

Three different scoring designs were developed to meet the needs of countries with varying numbers of respondents taking the paper-based instruments. These designs ensured a balanced distribution of bundles, or groups of booklets, across the number of scorers in a country while also balancing the order in which the bundles were scored. The Consortium also worked with countries that needed to deviate from these standard scoring designs, developing a tailored design to meet the country’s circumstances while still adhering to technical requirements.

Within each scoring design, of the following conditions had to be met:

- A minimum of 600 booklets sets (i.e., the set of booklets completed by a respondent) was required to be double scored using a balanced design to assess within-country scoring reliability. For some countries this meant that all booklets had to be double scored. Countries that collected more than 600 booklets had the option of single scoring booklets once the threshold of 600 was reached. For countries that collected fewer than 600 booklets, the guidelines required that 100% of the available booklets be double scored.

- Each scorer needed to score at least 125 items that were also scored by another scorer. This condition was necessary in order to generate enough data to evaluate the accuracy of the scorers.

- Two scorers were required to score the anchor booklets as specified in the scoring design to assess cross-country scoring reliability.

12.3.1 'Standard' three-scorer design

The standard three-scorer design was the default recommended design and applied to most participating countries. The design could be used in cases where countries collected *a total of around 600 booklet sets*. In this design, countries double scored all of their paper booklets, except for any extra bundles that were organized after this process was completed for the initial booklets. This design is presented in Table 12.1 below. Note that the numbers 1 and 2 shown in the table represent main (1) and reliability (2) scoring and not the scoring order. The design is summarized as follows:

- 18 bundles were assembled including:
  - C01 to C06 (Core Booklets)
  - L01 to L06 (Literacy Exercise Booklets 1), and
  - N01 to N06 (Numeracy Exercise Booklets 2).

  Within each booklet type, bundles included approximately equal numbers of booklets.

- Three bundles of anchor booklets were included, with 60 booklets in each bundle. As shown by the yellow highlighting, anchor bundle C00 included Core booklets, L00 included Exercise 1 booklets, and N00 Exercise 2 booklets. Each of these booklets was single scored.

- Three bundles (E01, E02 and E03) were reserved for any extra national paper booklets received after the initial booklet organization, bundling and dispersion took place. These booklets were single scored.

Table 12.1: Scoring design with three scorers

Bundle	Scorers		
	A	B	C
C01	1	2	
C02	2		1
C03		1	2
C00	1	2	
C04	2	1	
C05	1		2
C06		2	1
L01	1	2	
L02	2		1
L03		1	2
L00	1	2	
L04	2	1	
L05	1		2
L06		2	1
N01	1	2	
N02	2		1
N03		1	2
N00	1	2	
N04	2	1	
N05	1		2
N06		2	1
E01	1		
E02		1	
E03			1

As required, this design ensured that all scorers had a minimum of 125 scored items that could be matched to scores from other scorers.

The design required Scorers A and B to be bilingual as they scored the English language anchor booklets in bundles C00, L00 and N00.

12.3.2 Three-scorer design with single score bundles

If a country had *more than 600 booklet sets,* it could opt to use one of two scoring designs. It could use the standard three-scorer design described above and double score all of its Core Booklets, Exercise Booklets 1 and Exercise Booklets 2. It could also use the three-scorer design with single-score bundles presented in Table 12.2. In this design, 600 booklet sets were double scored to fulfill the requirements for the within-country reliability study, and the remaining were single scored. The three-scorer design with single score bundles is summarized as follows:

- As with the standard three-scorer design, 18 bundles were assembled including:
    - C01 to C06 (Core Booklets)
    - L01 to L06 (Literacy Exercise Booklets 1), and
    - N01 to N06 (Numeracy Exercise Booklets 2).

  These bundles included the 600 booklet sets to be double scored. Within each booklet type, bundles included approximately equal numbers of booklets.

- Additionally, nine bundles of national paper booklets were single scored. Bundles S01 to S03 were Core Booklets, S04 to S06 were Exercise Booklets 1, and S07 to S09 were Exercise Booklets 2. These bundles included the booklets remaining after the required 600 booklets were assembled for double scoring.

- Three bundles of anchor booklets were included, with 60 booklets in each bundle. As shown by the yellow highlighting, anchor bundle C00 included Core booklets, L00 included Exercise 1 booklets, and N00 included Exercise 2 booklets.

- Three bundles (E01, E02 and E03) were reserved for any extra national paper booklets received after the initial booklet organization, bundling and dispersion took place. Each of these booklets was single scored.

This design also ensured that all scorers had a minimum of 125 scored items that could be matched to scores from other scorers.

Table 12.2: Scoring design with three scorers and single score bundles

Bundle	Scorers		
	A	B	C
C01	1	2	
C02	2		1
C03		1	2
C00	1	2	
S01	1		
S02		1	
S03			1
C04	2	1	
C05	1		2
C06		2	1
L01	1	2	
L02	2		1
L03		1	2
L00	1	2	
S04	1		
S05		1	
S06			1
L04	2	1	
L05	1		2
L06		2	1
N01	1	2	
N02	2		1
N03		1	2
N00	1	2	
S07	1		
S08		1	
S09			1
N04	2	1	
N05	1		2
N06		2	1
E01	1		
E02		1	
E03			1

The design required Scorers A and B to be bilingual as they scored the English language anchor booklets in bundles C00, L00, and N00.

12.3.3 Two-scorer design

Although one of the three-scorer designs was appropriate for most countries, an alternative two-scorer design was also provided. This two-scorer design was used by countries that had 250 or fewer total booklet sets. The design ensured that each scorer would score at least 125 each of Exercise Booklet 1 and Exercise Booklet 2 as specified in guideline 11.3.2B in the *PIAAC Technical Standards and Guidelines*. The design is shown below in Table 12.3. As with the previous designs, note that the numbers 1 and 2 shown in the table represent main (1) and reliability (2) scoring and not the scoring order. The design was identical to the standard design for three scorers except that:

- Only one bundle, E01, was reserved for any extra national booklets received after the initial booklet organization, bundling and dispersion took place.

- Both scorers needed to be bilingual as they scored the English language anchor booklets in bundles C00, L00, and N00.

Table 12.3: Scoring design with two scorers

	A	B
C01	1	2
C02	2	1
C00	1	2
C03	2	1
C04	1	2
C05	1	2
C06	2	1
L01	1	2
L02	2	1
L00	1	2
L03	2	1
L04	1	2
L05	1	2
L06	2	1
N01	1	2
N02	2	1
N00	1	2
N03	2	1
N04	1	2
N05	1	2
N06	2	1
E01	1	2

In the Main Study, countries did not know, and could not control, how many respondents would take the paper instruments, as that was defined by the number of respondents who had no computer experience or failed the test of basic computer skills. Therefore, the Consortium recommended the following procedure:

1. Estimate the number of respondents who may go to the paper branch because they either did not have computer experience or failed the test of basic computer skills required for the assessment. This initial estimate was needed because countries selected the design they expected to use prior to scorer training.

2. Gather all returned national paper booklets, record their IDs in the appropriate tracking system, assemble and count the number of booklet sets.

   a) If the number of booklet sets is fewer than or equal to 250, use the two-scorer design.

   b) If the number of booklet sets is between 250 and 600, use the three-scorer design and double score every booklet set.

   c) If the number of booklet sets is greater than or equal to 600, choose one of these two options:

      - Option 1**:** double score all booklet sets using the three-scorer design.

- Option 2: use the three-scorer design with single score bundles, where a portion of the booklets are double scored for the reliability study and the remaining booklets are single scored.

  Options 1 and 2 were contingent on following these two rules:

1) *Rule 1: A minimum of 600 booklet sets must be double scored and used in the within-country reliability study.*

2) *Rule 2: Each scorer must have a minimum of 125 scores that can be matched to scores from one other scorer.*

12.4 Outcomes of the scoring reliability studies

Within- and cross-country reliability studies were conducted in both the PIAAC Field Test and Main Study.

The Main Study data showed a high degree of agreement for within-country scoring reliability, averaging 99.1% and surpassing the goal of 95%. It should be noted that a few countries showed 100% agreement between the main score and reliability score for one or more domains. This level of agreement has not been seen in previous international surveys of adult skills such as IALS and ALL. The most likely explanation for this finding is that in a few cases, countries implemented a resolution process that eliminated any scoring discrepancies.

The Main Study data also showed that average scoring accuracy across countries was very high, averaging 96.7% agreement. The cross-country reliability measures obtained from the anchor booklet scoring ranged from 89.9% to 98.5% across participating countries. Only three countries were below 95%. Thus the use of the anchor booklets verified that overall agreement across countries was good and allowed us to achieve common item parameters across countries, with very few items being assigned unique item parameters.

These data for both the within- and cross-country reliability studies demonstrate the success of international scoring training and the national application of that training. Overall, the data support that the result of this work by the Consortium and participating countries resulted in accurate and comparable scoring of the PIAAC paper-based items.

Chapter 13: Data Management Procedures

Ralph Carstens and Tim Daniel, IEA Data Processing and Research Center

13.1 Overview

In PIAAC, as in any multinational survey, it is a challenge to minimize total survey error, part of which can be introduced during capture, coding and processing of data. Subsequent steps in a survey process depend on the quality of the data that was originally collected. Errors during data capture, coding and processing of the data are difficult if not impossible from which to recover.

PIAAC administered an assessment of adult skills in two modes (computer and/or paper) in addition to a computer-assisted administration of a BQ. Design, data structures and formats in PIAAC are quite complex. For example, rich auxiliary and behavioral data, such as response times and navigation information were collected and processed in addition to the raw responses to support instrument validation, analysis and reporting.

Given these complexities – the timeline under which PIAAC was carried out and the diversity of contexts in which it was administered – it was imperative to standardize, as much as practically possible, the procedures as they relate to the national and international data management. A comprehensive manual, training sessions, a range of other materials, and in particular, a mandatory data management software were designed to help NPMs and their National Data Managers (NDMs; more on this role later) to carry out their tasks, prevent introduction of errors, and reduce the amount of effort and time involved in resolving them. Approaches had to be generally strict yet flexible at the same time to accommodate for some idiosyncrasies and needs (e.g., with respect to data sharing constraints) as part of the country-by-country data management process. In order to prepare a high-quality database (i.e., one that is valid, reliable and comparable) with the highest possible analytical utility, a variety of quality control processes and procedures were implemented.

This chapter summarizes the collaborative efforts, strategies and processes resulting in the rich, standardized international master database supporting all PIAAC reporting. The final PIAAC international master database included more than 1,700 international variables. In addition, more than 1,200 national variables (relating to adaptations and extensions) were defined, collected and processed for the 24 countries participating in the first round of the first cycle of PIAAC.

13.1.1 Tasks and responsibilities at the international level

The design and implementation of PIAAC was the responsibility of an international consortium of institutions led by Educational Testing Service (ETS). In this Consortium, the International Association for the Evaluation of Educational Achievement (IEA) Data Processing and Research Center (DPC) in Hamburg, Germany, had primary responsibility for designing, facilitating and supporting the data management at the national level, as well as the overall data management at the international level. In particular, the IEA DPC:

- proposed standards, guidelines and recommendations for the data work in countries;
- created and provided software, codebooks and manuals to countries;
- organized and conducted data management trainings;
- supported countries during the national database building;
- managed, processed and cleaned data at the international level;
- prepared analysis and dissemination databases for use by the Consortium, the OECD and countries; and
- provided data analysis software (see Chapter 23).

Conducting a study like PIAAC would not be possible without close cooperation and consultation among all stakeholders. These were the roles fulfilled by each partner in achieving a quality data product:

- ETS: review, cleaning, quality control and support with respect to interview workflow as well as cognitive response and log data (aggregate and full), release of data products to the Consortium, countries and the OECD;
- ROA: review, cleaning, quality control and support with respect to BQ data, questionnaire flow, harmonization of information from national adaptations, and coding of occupation and industry;
- Westat: review, cleaning, quality control and support with respect to sampling, weighting and survey operations related data; and
- OECD: overall review, quality control and support with respect to the resulting data products.

13.1.2 Tasks and responsibilities at the national level

Each participating country appointed an NPM to take responsibility for implementing PIAAC at the national level. The NPM had overall responsibility for ensuring that all required tasks, especially those relating to the production of a quality national database, were carried out on schedule and in accordance with the specified international standards and quality targets. The NPM was responsible for supervising, organizing and delegating all data management work. By "data management," we refer to the collective set of activities and tasks that each country had to perform to produce the required national database. This included the adaptation of codebooks, integration of data from the national PIAAC interview systems, manual capture of data after scoring, export/import of data required for coding (e.g., occupation), data verification and validation, and eventually submission of the national PIAAC database to the Consortium.

Because data-related tasks tend to be highly technical and require special skills, the Consortium recommended that an NDM be appointed by each NPM. The NDM was responsible for the day-to-day data management tasks within the country, was expected to carefully review all provided information and instructions, participate in all applicable trainings, supervise local data work, and, most importantly, communicate on data cleaning with the IEA in a timely manner. The NPM and NDM were expected to be supported by staff or organizations for manual data capture, scoring and coding during the applicable phases of the survey. The contribution that national technical personnel made was crucial to the survey's success and quality.

13.2 Key data management systems and integration processes at the National Center

13.2.1 Data management software, manuals and training

To standardize the national data work, countries were provided with a customized and extended version of the IEA Data Management Expert (DME) software originally designed and implemented for IEA work including Trends in International Mathematics and Science Study and Progress in International Reading Literacy Study. The IEA DME software supported data management at the National Center(s) after data collection. The IEA DME software was written in C# against the Microsoft .NET 4.0 framework and made use of a desktop, in-process variation of Microsoft SQL Server, more specifically, Microsoft SQL Server Compact 3.5 (SP2). Two versions of the software were created, one for the Field Test and one reflecting the revised assessment design and processes for the Main Study. The following list presents the key features of the IEA DME software and the customization to the PIAAC context:

- maintenance of a single, multi-table, robust and relational database for the integration of all sampling, response, workflow, log, scoring and coding data;
- documentation of the international as well as national record layout (codebook/code plan) and support for the addition and adaptation of national variables within constraints;
- extraction, transformation and storage of data from the various sources in PIAAC, most importantly the interview system;
- export and import to and from Excel; comma-separated and flat text files to interface with external processes, for example, the coding of occupation or the import of sample design data;
- manual data capture from scoring and response capture sheets as well as checks for double captured data;
- validation during import, manual entry and on demand by using pre-specified validation rules by variable, across variables, and across data sources using validity reports and statistics;
- supports for work on separate computers for data capture via file merging; and
- access control by using "roles" for managers and named data capture staff.

In concert with the IEA DME software, countries were provided with a comprehensive, 200-page data management manual detailing the processes, steps and checklists to be followed from the moment that the national interview systems, case management systems and paper instruments were readied for collection until the moment when national databases were submitted and initial data cleaning completed. Again, a Field Test manual and a revised/extended Main Study manual were provided.

Prior to the Field Test and again before the Main Study, NDMs or (in their absence) the NPM were expected to participate in comprehensive, data management trainings. Participation in these trainings was vital for the success of the project. These trainings focused on the setup and use of the provided

IEA DME software, the way it interacted with the assessment designs and interview system, the incorporation of national adaptations made in the BQ in codebooks, the integration testing between the national interview system and the data extraction logic, the import/export of relevant data stored in national case management systems or resulting from scoring processes, manual data capture from scoring sheets and the overall validation and verification of the database's completeness and consistency.

13.2.2 Codebook, database structure, record and value representation

Given the study's design and the technologies, the data structures and formats were relatively complex and somewhat different from those found in other school based large-scale international surveys. A variety of data sources were combined to build the national and international analysis and dissemination databases in PIAAC. The information in the database originated from the following assessment components, modules, sources and processes, mainly:

- sample design information (e.g., ID numbers, selection probabilities, stratification);
- screening and/or disposition information from countries' case management systems;
- interviewers' input into, or automatic import of, data into the case initialization module;
- interviewers' input into the BQ via the CAPI;
- behavioral/auxiliary information for the BQ (e.g., answers selection, timing, language changes, forward or backward navigation, consistency edits);
- interviewers' input and respondents' actions in the core modules;
- respondents' answers, detailed actions, timing and auto-assigned scores in the CBA;
- workflow information such as random numbers used in routing, automatically or interviewer assigned disposition codes, and timing information;
- respondents' original answers in the paper-based exercise and the reading components;
- countries' scoring and capture of scoring sheets for the paper-based exercise and the reading components (where used); and
- countries' coding of responses relating to the industry, occupation, language, country and region.

The PIAAC database included information from the sources above, and there was much more to consider. The interviews and exercises may have followed a variety of trajectories, data may have been generated for some respondents yet not others, and some data were captured during administration whereas other data were integrated after collection (for example, codes for occupation). Taking all this into account, the Consortium organized the data into a single relational database, though in multiple tables within this database. Each table corresponded to one or more modules in the survey. Later during the international data processing, most of these sources and tables were combined to form a more familiar "flat" analysis file.

The key concepts used in the PIAAC data management and database structure were *variables* (including their *value* and *missing schemes)*, *datasets* and *instruments*. The combinations of information in these entities form the PIAAC *codebooks* (elsewhere called *metadata* or *record layout*). In addition, data in PIAAC is stored by means of *data records* and eventually *data values.*

Variables correspond to fields (columns) in the resulting database. Each variable in PIAAC was defined by a set of attributes. The IEA DME software "reused" variable definitions in a number of ways. Variables were defined once, and only once, and then referenced in the corresponding *datasets* or *instruments* in which they were assembled. Secondly, *value* and *missing schemes* in the IEA DME software were defined only once and then referenced by the corresponding variables rather than being defined multiple times. This recycling of variables and schemes allowed efficient and consistent definition and adaptation of codebooks. *Variable attributes* were defined with the two most commonly used packages for statistical data, SAS and SPSS. Systematic and consistent *variable naming conventions* were applied for each component of PIAAC. Whereas variables of the BQ followed a naming convention derived from work at Statistics Canada, naming conventions for other assessment components followed a generic logic designed for PIAAC and took trend aspects into account (e.g., item naming found in IALs and ALL). Note, that variable names present in the exported interview system result files used a different naming convention and had to be renamed on import into the IEA DME database and for further analysis.

Each of the 33 datasets in PIAAC comprised the information for specific parts of the survey. A *dataset* is a logical collection of rows and columns where each column represents a clearly defined variable identified by its unique name and each row corresponds to a record of valid or missing values collected for a case or sampled person. Table 13.1 below describes the type of information they held along with the respective sources. Note that not all information was stored as part of the country database. Full cognitive log information was stored in its native format (XML) and provided to the Consortium at the time of data submission outside of the database maintained by the IEA DME software.

Instruments as used in the IEA DME software and database are logical sets of variables, i.e., a subset of variables selected in a particular sequence from a larger set of variables. Instruments were used for the manual data capture of paper scoring and response capture sheets.

Data records in the IEA DME software and database simply corresponded to a single row in a dataset, identified by one or more unique identifiers. Depending on a sampled person's path through the interview, data records for a single person existed in multiple but not all datasets. Each data record in a dataset had the same set of variables, and for each of these variables, either a valid value or a missing value was stored.

Table 13.1: Main Study datasets and sources

Dataset	Description	Specifics	Unique Identifiers	Source
SDIF	Sample Design International File	n/a	CASEID and/or PERSID (depending on sample design)	Imported from a country's study management system
BQR	BQ and global workflow	Results	PERSID	Extracted from BQ result files (XML)
BQL		Log	PERSID and SEQUENCE	Extracted from BQ log result files (XML)
BQC		Coded responses	PERSID	Imported from a country's coding process/system
CBR	Computer-based exercise	Results	PERSID	Extracted from cognitive result files (XML)
PCM1/ACM1	Paper Core Booklet (respondents or anchor)	**Main** scoring **First** capture	PERSID Secondary IDs: SCORERID_PPC, BOOKID_PPC, BUNDLEID_PPC, KEYOPID_PPC	Manually captured from core booklet scoring sheets
PCM2/ACM2		**Main** scoring **Second** capture		
PCR1/ACR1		**Reliability** scoring **First** capture		
PCR2/ACR2		**Reliability** scoring **Second** capture		
PLM1/ALM1	Paper Literacy Booklet (respondents or anchor)	**Main** scoring **First** capture	PERSID Secondary IDs: SCORERID_PP1, BOOKID_PP1, BUNDLEID_PP1, KEYOPID_PP1	Manually captured from literacy booklet scoring sheets
PLM2/ALM2		**Main** scoring **Second** capture		
PLR1/ALR1		**Reliability** scoring **First** capture		
PLR2/ALR2		**Reliability** scoring **Second** capture		
PNM1/ANM1	Paper Numeracy Booklet (respondents or anchor)	**Main** scoring **First** capture	PERSID Secondary IDs: SCORERID_PP2, BOOKID_PP2, BUNDLEID_PP2, KEYOPID_PP2	Manually captured from numeracy booklet scoring sheets
PNM2/ANM2		**Main** scoring **Second** capture		
PNR1/ANR1		**Reliability** scoring **First** capture		
PNR2/ANR2		**Reliability** scoring **Second** capture		
RCM1	Paper Reading Components Booklet	**Main** scoring **First** capture	PERSID Secondary IDs: SCORERID_PRC, BOOKID_PRC, BUNDLEID_PRC, KEYOPID_PRC	Manually captured from reading components response capture sheets
RCM2		**Main** scoring **Second** capture		
RCR1		**Reliability** scoring **First** capture		
RCR2		**Reliability** scoring **Second** capture		

Each logical dataset corresponded to a physical table in the relational database managed by the IEA DME software and had one or more identification variables in its first positions. Identification variables corresponded to units, entities or people in the survey or those that participated in its conduct. The identification variables used in the PIAAC Main Study are described below.

- CNTRYID: The country ID holds a 3-digit numeric code that follows the ISO 3166/UN M49 standard.

- CASEID: This is the household operational ID. It was assigned at the sampling stage for countries using a household sampling design.

- PERSID: This is the sampled person's operational identification number that uniquely identifies him or her. The PERSID variable appeared in all datasets as assigned at the sampling stage. In the case of household sampling, the PERSID was only assigned when within-household screening was completed and persons were sampled. The PERSID included a mandatory check digit based on approaches for universal product codes (UPC). The check digit proved to be highly efficient and effective in avoiding or identifying the vast majority of key entry mistakes.

- SCORERID_xxx: This ID identified the persons who scored paper-based exercise booklets on the corresponding sheets.

- KEYOPID_xxx: This ID identified the persons entering the values from scoring and/or response capture sheets, the key operators.

- BOOKID_xxx: PIAAC required countries to assign a unique booklet ID (serial number) to each printed paper-based exercise and reading component booklet.

- BUNDLEID_xxx: The bundle ID identified the bundles and their contained paper-based exercise booklets as defined by the international scoring design.

The following list provides a brief description of these datasets and the types of information they held:

- SDIF – Sample Design International File

    - The SDIF dataset held the required and optional variables as defined by the international sampling standards and included unique identifiers, sampling IDs, selection probabilities, stratification information, screening information, demographic information, disposition codes, information for variance estimation, raking dimensions and nonresponse adjustments variables.

- BQR – BQ and global workflow – Data

    - The dataset comprised explicit, implicit or derived variables captured as part of the general workflow, more specifically from the following case initialization module, the BQ (the bulk of the BQR dataset, hence the name), the CBA Core Stage 1 (ICT tasks), the administration of paper-based booklets (core, literacy, numeracy and reading components) and the observation module. The BQ variables in this dataset were subject to adaptation and extension, and any deviations from the international codebooks had to be reflected prior to production use.

- BQL – BQ and global workflow – Log

    - The interview system maintained a log file of actions and events relating to the same modules as described above for the BQR dataset. This log/audit dataset held information about the interviewer's actions during the CAPI, that is, any actions or variables that were explicitly shown on screen. This dataset contained multiple records per person. Each data record included information about the type of event, a timestamp, the item ID where the event occurred, and, where applicable, a value associated with the event depending on the type.

- BQC – BQ – Coded responses

    - Some of the answers to the BQ that were captured during the interview were subject to coding according to schemes for occupation (International Standard Classification of Occupations, or ISCO, 2008), industry (ISIC rev 4), language (ISO 639-2 alpha-3), country (UN M49 numeric) and region (TL2 OECD classification of geographical regions).

- CBR – Computer-based exercise – Results

    - The variables in this dataset represented the different pieces of information directly captured or derived from the computer-based exercise. It held all variables that were related to the computer-based literacy, numeracy and problem-solving items, more specifically the actual response; interim, and/or final scored responses; the number of defined action; the time elapsed before the respondent's first action; and the total time taken for the item.

- PCM1, PCM2, PCR1 and PCR2 – Paper Core Booklet

    - The PCM1 and the three related reliability ("R") and double capture ("2") datasets PCM2, PCR1 and PCR2 contained data for all items in the Paper Core Booklet. The responses to this booklet were scored on Core Booklet Scoring Sheets, and scored responses were captured and stored rather than the actual responses.

- PLM1, PLM2, PLR1, PLR2, PNM1, PNM2, PNR1 and PNR2 – Paper Literacy/Numeracy Booklet

    - The PLM1/PNM1 and the three related reliability ("R") and double-punching ("2") datasets PLM2/PNM2, PLR1/PNR1 and PLR2/PNR2 contained variables for all items in the Paper Literacy Booklet. The responses to this booklet were scored on Literacy Booklet Scoring Sheets and scored responses were stored rather than the actual responses.

- RCM1, RCM2, RCR1 and RCR2 – Paper Reading Components Booklet

    - The RCM1 and the three related reliability ("R") and double-punching ("2") datasets RCM2, RCR1 and RCR2 contained variables for all items in the Paper Reading Components Booklet. The responses to this booklet were captured on Reading Components Response Capture Sheets and, in contrast to the other paper-based booklets, actual responses were stored rather than the scored responses.

- AxM1, AxM2, AxR1 and AxR2 – Anchor booklets
    - These datasets held data originating from the anchor booklets scoring process in the cross-country scoring reliability study.

Each of the above datasets included records per person or case depending on the trajectory through the assessment. Each intersection of a variable and a record in the above datasets either held a valid or a missing value. Valid values were the individual pieces of collected information conforming to the corresponding variable specification, that is, the defined lengths, value schemes or ranges. The majority of variables in PIAAC were numeric and had a value scheme assigned to them (e.g., “1” corresponded to “Yes,” “2” corresponded to “No”). Wherever possible, value schemes limited the possible values that a variable could take. Missing data/values in a survey may occur when there are no data whatsoever for a respondent (unit nonresponse) or when some variables for a respondent are unknown, cannot be known, refused or otherwise not useful (item nonresponse). Missing data were distinguished semantically in essentially two broad groups: i) data that were missing by design, and ii) data that were supposed to be there but were not provided, or omitted. While missing data are inevitable in survey research, it is important to describe it properly and use it as information in itself to evaluate procedures, refine instruments or make assumptions about the mechanisms responsible as well as the likely consequences for the validity and possible bias of estimates. Analysis of item nonresponse is an important part of quality control, and consistent use of missing values ensured that the PIAAC data files contain detailed enough information on unit and item nonresponse (see also Chapter 16 on item-level nonresponse bias analysis).

The schemes to describe missing data in PIAAC during the time of data capture and building the national database were relatively simple and distinguished only a few types of missing data. In the following, the key missing value schemes used in PIAAC at the time of data integration are listed. A description of the missing values in the resulting public-use data products is presented in Chapter 23.

- Default missing scheme
    - This scheme was used for a large number of variables in PIAAC for which either a valid value was expected to exist for each and every data record or where there was no need to distinguish reasons for missing data during capture and database building.

- BQ missing scheme (numeric variables only)
    - All questions directed to the sampled person in the BQ explicitly included the options “refused” and “don’t know.” This missing scheme therefore distinguished the nature of the missing data and retained the information captured during the interview. The scheme applied to all BQ variables as well as most coded variables:
        - *Don’t know*: The sampled person was responsible for this type of item nonresponse by indicating “I don't know” or similarly.
        - *Refused*: The sampled person was responsible for this type of item nonresponse by refusing to answer the question.
        - *Not stated /inferred*: This is a systemic, catch-all nonresponse and was assigned if a variable was expected to hold a valid value but the value was missing, out of range, otherwise useless, and could not be reconciled or fixed.

- Free-text entry (FTE) missing scheme
    - PIAAC used a number of free text entry responses for occupation, field of industry, country, language, foreign qualifications and some other fields in the BQ. In order to retain the information provided by the respondents and/or the interviewer for later analysis and disambiguation, the IEA DME software imported missing values for any free text entries in the CAPI system as string constants, that is, either "[REFUSED]" or "[DON'T KNOW]".

13.2.3 National adaptations and extensions

Along with the IEA DME setup, countries were provided with an international codebook template. The international codebook for PIAAC included each and every variable and dataset known to the survey, including all variables relating to two international options (problem solving in technology-rich environments and reading components). Regardless of a country's participation in these options, the codebook and databases always included and displayed the variables for all components, but they simply had missing data if an option was not used.

The general approach to national variables was to include all international as well as all national data in a country's database in order to harmonize and map data post-collection. To do so, the international master codebook had to be adapted to reflect the national BQ in which countries adapted certain variables to their national and cultural settings as well as introducing additional national questions or adaptations/additions to the internationally ones. All adaptations and extensions applied in the national BQ had to be reflected in the codebooks as well in order to parse and store the information stored by the interview system. These adaptations related to the creation and specification of national variables, associated value schemes, as well as the adaptation of valid ranges for international variables as applicable (e.g., for currency units).

The adaptation of the international codebooks to reflect the national BQ was the responsibility of the NDM and performed according to instructions and guidelines provided by the Consortium. The international codebook template was used by NDMs as the starting point to which adaptations and extensions for national use were applied through controlled interfaces in the IEA DME software. The key input for this work was the national BQ itself as well as the agreed-upon Background Questionnaire Adaptation Spreadsheet (BQAS). As a key strategy, any adaptations to BQ questions had to be reflected under a new national variable name in order to clearly identify the likely need to harmonize, map or recode national to international variables after collection. A naming convention was applied that uniquely identified each national variable within and across countries. For example, a national variable for Germany that was based on item ABC would receive a name such as "ABC_DEU". In the case of extensions, that is, questions and variables unrelated to the domains and contents of the international BQ, a further variation using the suffix "X" would have been used, resulting for example in a name such as "ABC_DEUX". After all national adaptations were reflected in the codebooks, NDMs were responsible for thoroughly testing the import and correct mapping of data from the interview system, then submitting these codebook to the Consortium for further review, verification and for preparing international processing.

13.2.4 Data extraction from the computer-based delivery system

All data collected for PIAAC was integrated into a single national database managed by the IEA DME software. The primary means of integrating the database were by i) importing data from the national interview system, ii) manually entering the data via the data capture interfaces, or iii) importing data from national systems or processes. The bulk of the data in PIAAC naturally originated from the interview sessions and was stored in per-respondent result files in .zip format, each including a sizable set of XML format files for the various components of the assessment (BQ, core cognitive modules, main cognitive modules and observation module).

The contents of the per-respondent result file archives were generally stored as single records and mapped to the variables defined in the BQR, BQL and CBR datasets introduced earlier. In doing so, data were extracted from the individual XML files stored by the interview system, transformed as necessary, and then loaded into the respective target datasets (tables). Result data for the BQ was stored in datasets BQR and log data in BQL; cognitive result file information were combined from multiple XML files to form a single record in dataset CBR. The transformation comprised the mostly one-to-one mapping of values yet changed the data type from the generally used string types in the interview system to numeric values in the target database. For example, originally stored string literals such as "01" were stored as a numeric value "1". Missing values were mapped as well, from string literal "DK" for "Don't know" to a numeric value depending on the length of the variables (code 7, 97, 997 and so on). A refused response ("RF" in the result files) was mapped to numeric code 8, 98, 998 and so on in the database.

Additional transformation logic was applied in the following contexts:

- For multiple-choice items allowing more than one response in the BQ, values stored under the same name in the result files were mapped to individual variables.
- For currency values in the BQ data, any currency symbols were stripped.
- For numeric values with decimal places, thousand separators were stripped.
- For the BQ and workflow log data, string literals for event types were mapped to a numeric value scheme. For example, the event type "INTERVIEW_START" was mapped to the labeled value "1" in the target dataset BQL.
- Relating to workflow information, timer values for the reading components were transformed from string values formatted as "minutes:seconds:tenths" (e.g., "1:59:9") to tenths of seconds.
- For cognitive results, a name-mapping table matched long result variable names that were idiosyncratic to the interview system or sometimes not fully compliant with the naming conventions to shorter names used throughout all subsequent data products and analysis, such as names limited to eight characters in length.

As far as possible, the extraction and transformation logic checked for the integrity of the result file archive. However, given that some respondents broke off the interview and technical problems occasionally occurred, result files were parsed in a positivistic way, meaning that contents of the archives were parsed, provided that the main BQ result file existed along with any other materials found in the archive. As described before, NDMs were responsible for testing the integration between the interview system and the IEA DME maintained national database to make sure that i) all files, variables and values were mapped as expected, and ii) all nationally adapted or extended variables in the interview system were also reflected in the national codebooks. Certain values in the result files were only of interest at the time of collection and were not parsed and stored in the national database. For the most part, this related to strings for dynamic texts or interim values stored for some routing logic.

The full information, native CBA log files holding information on respondents' work on the cognitive assessment items were not parsed and loaded into the database. Instead, these were merely extracted from the result file archives and stored in separate folders. Countries were requested to provide these log files to the Consortium for further processing together with their initial data submission.

13.2.5 Data capture from scoring sheets and double capture reliability

Data capture is the process by which data collected on paper (e.g., on questionnaires, scoring sheets, or administrative records) are put in a machine-readable form. This section provides a description of the default process in PIAAC, that is, the recording of scored responses on scoring and response capture sheets and the subsequent capture of this information by means of the IEA DME's data capture interfaces.

According to the PIAAC technical standards and guidelines, the scoring of the paper-based exercise booklets had to be done twice by two different scorers following a scoring design recommended by the Consortium. Further, the manual data capture of each scoring (literacy/numeracy) or response capture (reading components) sheet had to be done twice by different key operators. This approach, although labor-intensive, allowed for separate investigation of error or disagreement introduced by the scoring and the data capture processes. This requirement also held for the international scoring bundle (anchor booklets) used in the cross-country reliability study.

This general data capture process was documented in detail in the data management manual along with advice on how to recruit, train and supervise key operators as well as operational recommendations for logistics, forming batches of materials for data capture and batch header examples. The manual entry of data in the IEA DME software was restricted to valid and missing values as defined by the respective scoring guides for literacy, numeracy and reading component items, and these permissible definitions were reflected in codebooks. The header of each scoring or response capture sheet included: the respondent's ID, the booklet ID, the scorer ID, the bundle ID, the score run (main or reliability) and the date of scoring. The information on the response capture sheets was simple and straightforward, allowing for efficient capture of data from sheets using numeric key pads. Respondent IDs were validated on capture. Similarly, out-of-range values or undefined codes were flagged and data capture was not allowed to proceed without correction. Partial entry was not supported. Each sheet had to be captured completely or not at all.

The set of rules provided to NDMs and their key operator staff included the following key advice:

- All scoring and response capture sheets must be fully completed before data entry can start. This included the header information on each sheet. In case there was missing, conflicting, otherwise or inexplicit information on any sheet handed to a key operator, these sheets must be returned to scorers (or the scoring process in more general terms) in order to be scored and filled correctly. Scorers were advised to revisit the original paper material in doing so.

- Data must be entered exactly as values appeared on the sheet, that is, without any corrections, unjustified interpretations or imputation.

- Checks for data capture accuracy and reliability must be checked on a regular basis, that is, at least once a week during the data capture process. This allowed the NDM to detect general misconceptions about the data capture rules or systematic misconceptions, underperformance or high incidental error rates of individual staff members. In addition, the Consortium recommended that the NDM monitor the accuracy of data entry on a more frequent, preferably daily, basis during the beginning of the manual data capture work.

The IEA DPC required countries to double enter all scoring/response capture sheets twice followed by a full reconciliation of any identified discrepancies by consulting the original paper materials. Checks for the accuracy and reliability of this double capture were built into the IEA DME software. This component listed all instances of disagreeing data and further provided an overall percentage of discrepancies. This procedure allowed the NDM to resolve data discrepancies before submission and the Consortium to estimate the agreement between key operators as well as the overall reliability of the manual data capture.

No margins were set for the acceptable levels of disagreeing data as a result of double capture. The Consortiums expected the manual key data capture to be 100% accurate and NDMs to resolve all identified discrepancies by revisiting the original scoring or response capture sheets and correcting the concerned values. All countries complied with this requirement and the evidence of data capture reliability provided by countries suggested that data were virtually free of data-capture error.

A number of countries requested permission to use alternative data capture means and processes. For example, some countries used scanning, followed by on-screen scoring processes, essentially collapsing the scoring and data capture processes into a single process. The Consortium carefully reviewed such plans and accepted deviations from the standard provided that countries were able to demonstrate similar or better quality. In these cases, the data resulting from these alternative processes were imported directly into the respective datasets.

13.2.6 Import of sample design data from study management systems

The SDIF was a mandatory deliverable from countries to the Consortium; the standard mode of transfer was as part of the national database. Countries were required to make use of one of the three supported import file formats (comma-separated, fixed length or Excel) to load SDIF-related data into the respective dataset. The actual import of the sample design information data into the SDIF dataset, using the import feature, was straightforward. Most variables in the SDIF were stored in a country's study management system. To import the sample design information countries had to:

- consult with Westat on the applicable variables in the SDIF to fill given the national sample design, plans for post-stratification and the like;
- export the applicable variables from the national study management system (or compiled/merged them from multiple data sources if applicable) into a single file in one of the import formats supported by the IEA DME software;
- ensure that the data contained were complete, accurate and formatted as defined by the respective codebook;
- ensure that variables not applicable to the national sample design were either represented by blanks in fixed-length import files and empty cells in CSV and Excel, or not included at all in the import file;
- ensure that all records in the import file were uniquely identified by a valid person ID and/or case ID as applicable; and
- ensure that any numerical variables used no more than the specified number of decimals.

Whereas the above stated prerequisites as well as file structure and variable definitions were automatically validated on import, no checks for completeness of SDIF data could be run given the varying sample designs across countries. Sampling- and weighting-related data were reviewed by Westat following the submission of national databases, and numerous corrections and additions were processed for a large number of countries until a complete, valid and accurate SDIF could be finalized and receive signoff prior to weighting.

13.2.7 Import of coding data from external processes

A number of free text entry variables in the BQ were not only captured during the interview but were subject to coding according to schemes for:

- Education: International Standard Classification of Education, or ISCED, 1997 long, ISCED 1997 broad fields of education and training, ISCED 1997 short
- Occupation: ISCO 2008 at the four-digit unit group level
- Industry: International Standard Industrial Classification of All Economic Activities, or ISIC, Revision 4 at the four-digit class level
- Language: ISO 639-2/T (alpha-3/terminologic)
- Country: UN M49 numeric
- Region: TL2 level of the OECD classification of geographical regions

The BQ variables that served as inputs for coding, as stipulated by the BQ framework as well as the corresponding validation and analysis plans, were documented as part of the data management manual. Related advice and training was given to countries as part of data management trainings. Separate expert trainings were held for the coding of occupation against the ISCO standard and industry against the ISIC standard. The respective coding schemes were included in the codebooks to facilitate validation at the time of database integration.

More specifically, the following instances of coding were required from countries:

- Coding/mapping general ISCED responses
    - All countries posed education-related questions in a closed format using national classification. In that sense, no actual coding was carried out (except in the case of “foreign qualifications” that had to be coded; see below). Countries either converted these national codes into ISCED 1997 themselves or provided conversion rules. Countries were required to deliver both the code in the national classification and the corresponding international code.
    - Countries were required to code the highest foreign qualification for all respondents who reported a foreign qualification using responses to B_S01a1, the name of the “foreign” highest qualification (write-in), and B_Q01a3, the nationally corresponding level of the “foreign” highest qualification (a nationally adapted list).
    - ISCED codes for respondents’ highest foreign qualification were stored in variable ISCED_HF in the BQC dataset.

- The missing scheme for the variable ISCED_HF was the standard numeric scheme for the BQ. Because ISECD_HF was of length 2, the missing codes were also of length 2:

    - *Don't know* was used if the two raw responses were marked as "don't know."
    - *Refused* was used if at least one raw response was marked as "refused."
    - *Not stated* was used if at least one raw response was given but not interpretable or otherwise useless and it could not be reconciled or fixed.

- Coding of occupation to ISCO 2008 and coding of industry to ISIC Revision 4

    - Four-digit codes from the 2008 ISCO-08 were used to code the occupation of the respondent (current and last job as applicable). The corresponding target variables in the BQC dataset were: ISCO08_C (current job) and ISCO08_L (last job).

    - Countries that opted to initially code in ISCO 1988 were made aware that no automatic conversion from the ISCO 1988 to ISCO 2008 existed: certain codes in ISCO 1988 were split up into multiple codes in ISCO 2008, while other codes were merged. Therefore a manual verification of the correspondence was required for these codes.

    - If a country had coded in ISCO 1988, this coding had to be provided as well as the coding in ISCO 2008. The corresponding target variables in the BQC dataset were: ISCO88_C (current job) and ISCO88_L (last job).

    - Four-digit codes from ISIC, Revision 4[1], were used to directly code the sector in which the respondent was working (current and last job as applicable). The corresponding target variables in the BQC dataset were: ISIC4_C (current job) and ISIC4_L (last job).

    - The missing scheme for the variables ISCO08_C, ISCO08_L, ISCO88_C, ISCO88_L, ISIC4_C and ISIC4_L was a special numeric scheme. Because the ISCO/ISIC variables were strings of length 4, the missing codes were also strings of length 4:

        - Don't know (code "9997") was used if all of the raw responses were marked as "don't know."
        - Refused (code "9998") was used if one or all of the raw responses were marked as "refused."
        - Not stated (code "9999") was used if at least one raw response was given but not interpretable or otherwise useless and it could not be reconciled or fixed.

    - The coding of occupation and industry to ISCO/ISIC was subject to quality control implemented by ROA. As part of the data submission, countries were required to provide corresponding evidence and reports comparing the unweighted and weighted distributions of occupational groups at the two-digit level to external information from, for example, the most recent national labor-force survey.

    - Responses that could not be coded at the four-digit level, that is, codes at the one-, two-, or three-digit level, were subjected to review by a coding expert.

[1] http://unstats.un.org/unsd/cr/registry/isic-4.asp

  - Some countries were not legally able to disclose ISCO/ISIC data at the four-digit level and submitted data only at the permissible level of detail.

- Coding of language to ISO 639-2/T

  - For language-related free-text entries, the ISO 639-2/T alpha3 (terminologic) scheme was used.

  - The corresponding target variables in the BQC dataset were LNG_L1 (first language learned at home and still understood), LNG_L2 (second language learned at home and still understood) and LNG_HOME (language most often spoken at home). By their very nature, ISO 639-2 three-digit alphanumeric codes for languages were defined as strings of length 3 in the BQC dataset.

  - The coding of languages involved two steps:

    - Mapping the numeric responses to the national closed format language questions in the BQ to the codes in ISO 639-2.
    - Coding the write-in responses to the "other" languages questions in the BQ to the codes in ISO 639-2.

  - The missing scheme for the variables LNG_L1, LNG_L2 and LNG_HOME was a special numeric scheme. Because the ISO 639-2 variables were strings of length 3, the missing codes were also strings of length 3:

    - Don't know (code "997") was used if the raw response was marked as "don't know."
    - Refused (code "998") was used if the raw response was marked as "refused."
    - Not stated (code "999") was used if a raw response was given but not interpretable, otherwise useless, not covered by the scheme and it could not be reconciled or fixed.

- Coding of country to UN M49

  - Countries coded the country names in various questions of the BQ using the numerical codes of UN M49. In most cases, a country-specific list of countries was used that covered the most relevant countries plus a category "other." Both the "listed" countries as well as the "other" category were converted by the countries into UN M49.

  - The name of the country reflected the CURRENT name of the country in which the highest qualification was attained or in which the respondent was born, not the name of the country in the past (regardless of whether the question related to the past, e.g., country of birth).

  - The corresponding target variables in the BQC dataset were CNT_H (country of highest qualification) and CNT_BRTH (country of birth). UN M49 country codes were defined as integers of length 3 in the BQC dataset.

  - The coding of countries involved two steps:

        - Mapping the numeric responses to the national closed format country questions in the BQ to the codes in UN M49.
        - Coding the write-in responses to the "other" country questions in the BQ to the codes in UN M49.
    - The missing scheme for the variables CNT_H and CNT_BRTH was the standard, numeric BQ missing scheme. For the coding of country, the missing codes were used as follows:
        - Don't know was used if the raw response was marked as "don't know."
        - Refused was used if the raw response was marked as "refused."
        - Not stated was used if a raw response was given but not interpretable, otherwise useless, not covered by the scheme and it could not be reconciled or fixed.
- Coding of region to OECD TL2
    - Countries were required to code the geographical region corresponding to the respondent's address at the TL2 level using the OECD classification of geographical regions, for example, "DE6" for a respondent in Hamburg, Germany. The corresponding target variable in the BQC dataset was REG_TL2, and this variable was defined as a string of length 5 in the BQC dataset.
    - The variable REG_TL2 was not derived from BQ responses but from sampling/survey control data. Therefore, the missing scheme for the variables REG_TL2 was the default missing scheme that only permitted "blank" as a missing value as data were expected to be available for all sampled persons.

13.3 Data verification and editing at the National Center

13.3.1 Validation, verification and record consistency checking prior to data submission

Each country was required to perform verification of the national database to identify and, if necessary, resolve errors and inconsistencies in the data. For carrying out this important part of the quality control work, tools to apply the minimally required checks as well as policies regarding the within-country editing of data were provided to countries.

Automated validation checks to detect values outside of the defined range for a variable, duplicate IDs and double data capture checks to detect and resolve data capture errors were made available as part of the IEA DME software. These checks were designed as an initial inspection of severe gaps or mismatches in the data and not intended to replace the more thorough data-cleaning process at the international level that was done centrally. Countries were required to run these on a regular basis. Further, record consistency checks were included in the software. The record consistency checks included 45 checks that identified possible inconsistent records across datasets. The checks were consecutively numbered and grouped by content:

- Checks 1 to 24 flagged possible linkage problems between datasets, that is, they listed respondent IDs that were expected to be found in a dataset, given their existence in another one and the interview flow;

- Checks 25 to 30 flagged possible problems in the scoring datasets, for example, an insufficient number of anchor booklets contained;
- Checks 31 to 43 flagged possible problems related to sampling information, for example, indication that two persons were sampled in the household but only one record for this household existed; and
- Checks 44 and 45 reported problems of general nature, especially related to technical problems and "out of design" cases.

In addition to the automated and consistency checks, the IEA DME software contained facilities to review descriptive statistics, including minimum, maximum, mean, median, variance, percentiles and sample standard deviations, as well as to cross tabulate variables for quality control purposes. NDMs were strongly urged to review frequency distributions of their data for plausibility and/or agreement with expectations. It was also important to verify the completeness and integrity of the database with respect to the included data records. Sampled persons in PIAAC followed a variety of paths through the interview, each generating records in one or more datasets yet not in others. In addition, the existence of data records also depended on whether the sampled person completed the entire interview, or broke off before its end and consequently didn't work on all of the applicable components. NDMs were advised and trained on the importance of checking the number and IDs of data records existing in the various tables of the database against the known and therefore expected numbers from survey records and study management systems.

13.3.2 Permissible and prohibited data editing and cleaning

Countries were requested to run the checks described so far in this chapter to ensure, as much as possible, that the within-country data capture and integration accurately and authentically reflected the values given by the sampled persons and/or the interviewers.

Countries were asked to refrain from implementing any type of general data-cleaning or data-flow editing on their own prior to the submission of the data. The Consortium partners requested original access to the types and the magnitude of, for example, outliers, implausible values or implausible combinations of these in order to refine the instruments and/or to identify problems with the translation of questionnaire items. However, countries were encouraged to make corrections to the data that were clearly attributable to the survey process, data-capture mistakes or similar misunderstandings made by, for example, the interviewer. Common examples of these edits included the correction of incorrectly recorded disposition codes or incorrect secondary IDs (e.g., booklet IDs). This was considered to be a part of the normal and mandated data verification and checking. Also, exceptions applied to instances of technical problems in the virtual machine (VM) where a disposition code "90" may have had to be assigned after data collection in those cases where on-site recovery was impossible and only partial data (or none at all) was extracted from the VM. Other exceptions related to reproducible and verified error sources, for example, residual BQ routing errors, recoding errors and so on which could be corrected using logical and verified correction procedures.

The Consortium received a number of requests to change/edit the data in order to make it more consistent across variables or more consistent with other data collections. The Consortium's consistent position communicated to countries was that data collected during the interview took precedence over wholesale interpretations or assumption without concrete verification or evidence indicating that originally corrected data were unreliable or invalid. Where no additional data collection was conducted, or counter information was available to override the original information, no change was implemented or allowed. Exceptions were related to reproducible errors (e.g., routing, recoding, etc.). A small number of verifiable exceptions were made but required written documentation and pre-approval by the Consortium.

13.3.3 Confidentiality review, editing and variable suppression

Some countries had regulations and laws in place that restricted the sharing of data, as originally collected, with the Consortium and/or the OECD. The key goal of such disclosure control is usually to prevent the spontaneous or intentional identification of individuals in the full-information microdata. On the other hand, suppression of information or reduction of detail clearly has an impact on the analytical utility of the data. Therefore, both goals had to be carefully balanced. As a general directive, the OECD requested all countries to make available the largest permissible set of information at the highest level of disaggregation possible.

A small number of directly identifying variables that were collected during the case initialization were suppressed by default in any database exported for submission to the Consortium. This included the respondent's name, address, and telephone number. According to the technical standards, each country had to provide the Consortium with early notification of any rules affecting the disclosure and sharing of PIAAC sampling, operational or response data. Furthermore, each country was responsible for implementing any additional confidentiality measures in the database before delivery to the Consortium. Countries especially reviewed the sample design information (dataset SDIF) and the variables collected through the BQ (dataset BQR) with respect to indirectly identifying variables or otherwise sensitive information. Most importantly, any confidentiality edits changing the response values had to be applied prior to submitting data to the Consortium in order to work with identical values during processing, cleaning and analysis. The IEA DME software only supported the suppression of entire variables. All other measures had to be implemented under the responsibility of the country via the export/import functionality or by editing individual data cells.

The Consortium asked for complete and detailed documentation about any implemented measures to evaluate the impact on the analytical utility of the dataset, especially with respect to the introduction of bias, attenuation of within-variable variance, or between-variable correlations as a result of data suppression or perturbation. The majority of countries suppressed data at the variable level and submitted a database excluding certain types of information such as birth countries, original free text entries, full log information or detailed earnings values. These suppressions were carried forward throughout all subsequent data processing and analysis stages and into the public-use data products. Perturbation of original values according to the documentation known to the Consortium applied in two instances:

- Austria used statistical coarsening for the original, detailed earnings values (micro-aggregation).

- The United States perturbed data prior to submission following local standard operating procedures for large-scale surveys. Within-record consistency was maintained. The Consortium received no detailed account of these perturbations and consequently was unable to review, validate or assess the impact of these edits on the data or any inferences based on it.

A general procedure for the suppression of information from the for public-use databases was implemented after processing. These additional suppressions were handled by the Consortium in a standardized way. Exceptions to the general rule of suppressing an entire variable apply in these cases:

- Austria and Estonia suppressed single values given small frequencies for some language and country variables.

- Canada applied a small number of case-level suppressions that held values or combinations believed to identify sample and or population uniques.

13.3.4 Data submission and required documentation

After the collection, integration and verification of data, each country was responsible for submitting the required materials to the Consortium. The materials to be submitted electronically to the Consortium after the Main Study were the following:

- A single, integrated, verified, confidential and exported database per country in the IEA DME's format using the adapted national codebooks, that is, including all national variables and values (except for suppressions).

- A single zip archive including all original cognitive log files extracted and stored as part of the data parsing from the interview system.

- A free-format documentation with double-coding reliability evidence and explanations for QC purposes according to the technical standards and guidelines. The information requested comprised tables in which countries compared data collected in PIAAC with the most recent labor force survey (or equivalent) on the distribution of i) highest level of education, ii) labor force status, ii) occupation at the one- and two-digit level (ISCO 2008), as well as iv) sector of industry in 21 sections (ISIC, A-U).

- A comprehensive and detailed free-format documentation of implemented confidentiality edits, if any, and the effect of these edits on univariate and multivariate properties.

- A comprehensive and detailed free-format documentation of any other issues or notes that required attention by the Consortium during data processing and analysis. The document was expected to include notes for example pertaining to out-of-design cases, that is, respondents that did not follow the assessment design as prescribed or technical problems.

On export from the IEA DME software, a copy of the current national database was created. All values for all occurrences of a variable marked as "suppressed" in the codebook were set to blank values in the exported database. The national database exported was marked as non-productive and read-only.

Any data submission to the Consortium had to be made through secure channels. For this purpose, a SSL/TLS secured FTP site and a corresponding Web interface were set up. Document exchange folders were created for each country. Access to such a country exchange folder was limited to authorized staff members of the Consortium and the national center.

13.4 Data processing and editing at the international level

This section describes the process from the moment that national databases were received from countries until the moment that a preliminary international database, consisting of each national database, was produced. The main international data processing phase stretched from June to October 2012. The initial phase (June-July 2012) was used to clean data at the case level and with respect to all relevant fields in order to prepare and flag cases for weighting that are valid and comply with the PIAAC definition of "complete." The following months (August through October 2012) were used for any residual data cleaning and/or for the processing of additional, revised or erroneous data. Exceptions to this general timeline apply given the slightly differing schedules in countries' data submissions.

In general, the data processing for PIAAC was straightforward, carried out separately for each country, yet based on a common framework of standardized procedures and edits applicable to all countries. The bulk of the data processing was implemented using SAS version 9.2. All data processing was run in Unicode mode, thereby preserving all national strings in free text entry variables. Programs for initiating and controlling SAS or other processing programs were based on generalized processing systems used across all IEA and third-party surveys managed by the IEA DPC. All processing systems were set up so that the different steps, from import to exporting data products, could be run again to include and reflect all changes and edits. Missing values were represented using SAS standard (".") or special missing values (".A"-".Z").

13.4.1 Data import and structural reorganization

The import and merge of data essentially followed the below sequence of steps. As a first step, data capture accuracy was checked using the submitted IEA DME database and recorded. As noted before, data capture accuracy was found to be satisfactory for all participating countries in the Main Study. Data from the double capture process were set aside and not processed further.

Next, each national database in the DME's native Microsoft SQL Server Compact format were loaded into a temporary Microsoft SQL Server 2008 R2 server database "as is," that is, without any transformations or changes. Using these SQL server data as the input, a SAS-based program read all data from the national databases, merged tables as necessary and checked for structural integrity and deviations from the international variable layout. This step produced four SAS formatted files.

Original national database tables were consecutively merged using PERSID to form a single flat file named PRG (for PIAAC Response General) encompassing all variables for a single case from the following source datasets (see Section 13.2.2 above for details):

- SDIF – Sample Design International File
- BQR – BQ results and workflow
- BQC – Coded responses
- CBR – Computer-based exercise results
- PCM1 – Paper core booklet results
- PLM1 – Paper literacy booklet results
- PNM1 – Paper numeracy booklet results
- RCM1 – Paper reading components results

Cases or respondents present in neither the SDIF nor BQR dataset were dropped at this stage. The PRG file was inclusive of all national adaptations and extensions introduced by countries.

The dataset in the national database holding reliability scoring/capture data were merged using PERSID to form a flat file named PRR (for PIAAC Response Reliability), encompassing all variables for a single case from the following source datasets:

- PCR1 – Paper core booklet results

- PLR1 – Paper literacy booklet results
- PNR1 – Paper numeracy booklet results
- RCR1 – Paper reading components capture results

The IEA DME dataset holding reliability scoring/capture data were merged using PERSID to form a flat file named PAG (for PIAAC Anchor General) encompassing all variables for a single case from the following source datasets:

- ACM1 – Anchor core booklet results
- ALM1 – Anchor literacy booklet results
- ANM1 – Anchor numeracy booklet results

The IEA DME dataset holding reliability scoring/capture data were merged using PERSID to form a flat file named PAR (for PIAAC Anchor Reliability) encompassing all variables for a single case from the following source datasets:

- ACR1 – Anchor core booklet results
- ALR1 – Anchor literacy booklet results
- ANR1 – Anchor numeracy booklet results

One additional file named PRL (for PIAAC Response Log) was produced from the information parsed in the national database's BQL dataset. This file was not subject to cleaning or editing as it mainly included timing information for validation purposes.

For each component and source table, a flag was created regarding whether data relating to the case existed in the source dataset with only missing values, with some valid values, or with a complete set of values.

13.4.2 Structure check and recoding of national adaptations

The structure check stage performed several checks that related to file and variable structure integrity. It checked for changes in international variable definitions, availability of mandatory variables applicable to all sample designs and contexts, as well as the validity of national variable definitions with respect to naming conventions and in light of agreed-upon adaptations in the BQAS. All original missing values in national databases were programmatically mapped to SAS missing values on import. At this stage, validation checks for all numerical variables ran and ascertained that no unconfirmed out-of-range values remained in the data. NDMs received standardized reports on any flagged inconsistencies for either confirmation or resolution.

Questions in the PIAAC master BQ were designed to have the same meaning for respondents in all participating countries irrespective of differences in language and culture. However, two sets of adaptations or extensions had to be applied by countries in the process of translation/adaptation: i) mandatory adaptations in the case of ISCED levels, country name placeholders, and the like, and ii) idiosyncratic adaptations and extensions that reflected national research interest or were used to align questions with other data collections. These national adaptations and extensions had to be processed along with data for not adapted questions. While national extensions were processed, returned to countries for their own use, and also referenced in the psychometric analysis, data collected from national adapted questions had to be harmonized by means of recoding for it to be internationally comparable.

For this purpose, the IEA DPC processed and reviewed all final BQAS and created Excel documents that only included national extensions and those structurally adapted (e.g., added response options). The result from this process was documentation of country adaptations requiring attention during the international data processing phase by recoding national responses back to the international response schemes and variables where needed. Additionally, it was recorded for each adaptation whether a recoding was needed and, if yes, whether the IEA DPC or the country was responsible for implementing it. These “reverse” BQAS sheets were discussed with the concerned country and finally reviewed by ROA, the Consortium partner initially responsible for reviewing and approving national adaptations.

The recodings due to national adaptations were applied by default during the course of processing countries’ data according to agreement found in the process described above. National variables affected by these adaptations retained their original values through the whole cleaning process and provided to countries unchanged after data processing. Many countries, though, decided to perform several, if not all, necessary recodings themselves prior to data submission. This was supported and approved by the Consortium in cases where countries also provided the constituent national variables referenced in the recodings. In some cases, countries used complex adaptations in the BQ, and this in turn resulted in very complex recodings that had to be harmonized under country responsibility and local validation and verification. In some other cases, countries were responsible for recoding data prior to submission given confidentiality reasons, that is, situations where countries were not able to release certain variables to the Consortium due to national legislation.

The Consortium reviewed the appropriateness of all applied recodings with respect to international comparability of data by means of cross tabulations using a single or multiple source and target variables. This also applied to cases where countries applied recodings prior to data submission and the source national variables were provided to the Consortium. For recodings where the original national variables were not disclosed to the Consortium, no detailed validation of the recoding process was possible and the Consortium informed the concerned countries that any error as a result of these recodings was entirely the responsibility of the country. Nonetheless, the Consortium applied coarse and technical plausibility checks of the resulting data. Countries were provided with the same frequency distributions in the resulting data and were asked to check and verify them. Table 13.2 provides an overview where recodings were applied and whether the national variables referenced were available to the Consortium. Following from the process descriptions provided by those countries which applied recodings prior to submission, the Consortium was not aware of any indication that particular recodings applied by countries were invalid or flawed in other ways. However, the volume of national questions and variables (in excess of 1,200 variables across the 24 countries), the complexity of some adaptations and extensions, a somewhat different response process, and differential missing data in cases where multiple questions were referenced to yield an international value made it quite likely that some minor errors remained undetected in the data.

Table 13.2: Responsibility for, and time of, mapping national to international variables

Country name	All mappings applied by country prior to submission	Some mappings applied by country prior to submission	All mappings applied by Consortium after submission	Consortium had no access to some or all original national data
Australia	X			X
Austria		X		
Canada	X			
Cyprus[2]			X	
Czech Republic	X			
Denmark			X	
England/N. Ireland (UK)	X			
Estonia		X		
Finland	X			
Flanders (Belgium)			X	
France			X	
Germany		X		
Ireland			X	
Italy	X			
Japan			X	
Korea			X	
Netherlands	X			
Norway	X			
Poland	X			
Russian Federation[3]			X	
Slovak Republic			X	
Spain			X	
Sweden	X			
United States			X	

13.4.3 Data cleaning process, systems, communication and reports

For the PIAAC Main Study, a comprehensive set of checks was implemented that allowed for a broad as well as deep inspection and cleaning of data files. As stated initially, this process of cleaning involved the Consortium partners directly involved in the database building, the OECD as the primary data consumer at the international level, and last but not least the NDMs and NPMs in each country.

[2] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[3] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

As part of the data cleaning process, records and variables were checked for consistency, that is, that no duplicate IDs existed, no unaccounted for wild codes existed, and the expected data pattern (given for example a case's booklet assignment, disposition codes or ICT core pass status) matched the observed data patterns. Additional checks focused on the consistency of records and variables within a dataset, the linkage of records between datasets, as well as repeating soft validation checks already run during the interview. Any flagged issues had to be reviewed, verified, resolved or, where this was not possible, at least commented on by countries. Extensive and detailed communication between the IEA DPC and any participating country on data inconsistencies and their resolution took place and detailed reports were provided to NPMs and their NDMs on any such issues, and they were asked for confirmation or advice.

The overall observed quality of the submitted data was usually very good. There were no substantial structural errors in databases and almost all cases matched between data sources. In total, only about a dozen or so cases out of more than 150,000 had to be removed or corrected because they were out of scope – for example, if they included both computer- and paper-based data where only one of the two was expected given the respondent's trajectory. The high degree of internal consistency of the data can probably be attributed to three main factors: i) the fact that the PIAAC assessment was highly standardized and computer-controlled and, technical problems aside, provided no possibility to follow an incorrect path, ii) the use of strict ID validation in all components of the survey, and iii) the diligent work of NDMs in identifying the few mismatching cases and allocating data as appropriate. Where data were not matching the expected design, narrative reports from countries indicated that this was due to interviewers not following the intended workflow. For example, some interviewers administered paper booklets in instances where there was a technical problem with the CBA portion of the assessment. Data values for components not applicable to a respondent were, after careful inspection, reset to their respective missing codes.

Other potential issues were mostly related to incidental, variable-level errors. These were too diverse and too sparse to be reported here in any detail. Recurring issues across countries included, but were not limited, to:

- incorrect or inconsistent disposition code assignment, often in cases of technical problems;
- incomplete data for anchor booklets used for cross-country reliability analysis;
- incomplete or incorrect mapping of national adaptations where this was the country's responsibility;
- missing indication of suppressed variables at the time of data submission;
- discrepancies between age and gender as recorded in sampling frames, collected via screeners, entered or loaded during case initialization, or reported by respondents in the field;
- incomplete loading of sample design data for noninterviews, that is, individuals who refused to take the interview, language-related nonresponse, or absences;
- incorrect loading of other sample design data given a country's plans for nonresponse adjustment and raking; and

- incorrect, incomplete or unreliable coding – for example, concurrent mapping and coding of country of birth responses, agreement of occupational distributions with external data sources, or handling of occupation/industry codes as numerical.

In many cases, such issues could be resolved by reviewing the original instruments and reviewing registry information and/or feedback from field operation, scoring or data entry staff.

13.4.4 General edits and the iterative integration of derived variables

After the cleaning phase, the state of the data reflected the data collection as accurately as possible, and any individual or structural inconsistencies were removed to the extent possible and known to the Consortium. During this post-cleaning phase, general structural changes were made to the data for all countries based on the now cleaned original values.

Most importantly, the processing systems reimplemented the routing logic included in the international master BQ and assigned "valid skip" missing values to any variable that was not in the respondent's path. It should be noted that this was done in a comparable way for all countries and only considered original or recoded responses to any international BQ variables. "Valid skip" missing codes were generally only assigned for respondents who started the BQ. In the case of breakoff during the BQ, the "valid skip" recoding was only implemented up until the last known valid value. All subsequent values were coded as "not stated." Further, "valid skip" codes were carried forward to any coded variable (e.g., second language learned) if the referenced original variables were previously coded as "valid skip."

In a few cases, countries not only adapted questions but were given permission to also adapt the routing rules implemented in the international master BQ. This resulted in a few instances where too little or too much information was collected in comparison to a route that a "standard" respondent would have taken through the BQ. Excess data collected due to national routing was overwritten with "valid skip" codes in the process described above for reasons of international comparability because respondents affected were not supposed to have data observed according to the international routing rules. In cases where too little information was collected, and thus missing data were present yet not expected, there was usually no way to recover from this. Such data cells were coded as "not stated." The overall number of affected cases was very small (a maximum of 207 cases with excess data and 38 cases with missing data for a few variables were present in one country) but nevertheless shows the risk and possible impact of excessive national adaptations to already complex international collection instruments.

Further at this stage, "not reached" codes were assigned to cognitive assessment items in the paper path. For this, items with value 0 = "no response" were recoded to "not reached/not attempted" according to a generic algorithm that checked for "no response" values from the end of each item block (and individually for each item block) and assigned value "not reached" until a valid code was encountered. "Not reached" codes were also assigned to item responses in the computer-based path. These adjustments were done at ETS, delivered to the IEA DPC as a set of mergeable files with revised data, and integrated into the master databases on each run.

This processing phase was further used to derive or merge reporting variables, weights and scale scores. The process of deriving variables was highly iterative and depended on the progress of the weighting and analysis. The derivation and integration observed the necessary sequencing conditions. For example, scripts for the coarsening of variables had to be based on revised original and/or derived variables.

- Derivation of variables from sample design and case initialization data
    - A number of sample-design related variables were derived from sample design, case initialization and BQ information. With the exception of some special settings in some countries, the derivation of these variables was done according to standardized scripts that were consistently applied with each pass of the data for all countries. These sampling-related variables were independently computed by IEA DPC and Westat and compared as well as reconciled as necessary. The most important derived variables in this segment were:
        - Three final, combined disposition codes for the case initialization and BQ phase (DISP_CIBQ), for the main assessment (DISP_MAIN) and including reading components for those countries participating in the option (DISP_MAINWRC).
        - Resolved age (AGE_R) and gender (GENDER_R) taking into account frame information but giving precedence to observed data during the interview, further incorporating collected age and gender in the case of literacy-related nonresponse.
        - A completion flag (COMPLETEFLG) set according to technical standards definitions in relation to assessment components and/or key items.
        - A weighting flag (WEIGHTFLG) computed from the disposition codes and/or literacy-related nonresponse information.
        - An interim code (SCENARIO) derived according to a set of rules intended to identify cases earmarked for weighting yet with insufficient information or vice versa.
    - The key Consortium partners responsible for identifying valid cases reviewed the outcomes of the above assignment in regular online meetings and revised the weighting and completion flags as well as aggregate disposition codes in a small number of cases depending on whether sufficient information was available to assign a weight and/or analyze the cases.

- Integration of weighting and variance estimation variables
    - Once valid cases where flagged for weighting and analysis, weights and scale scores were merged to countries' data files as they became available. Weights were computed by either the concerned countries or Westat and were merged to the files.

- Derivation of variables from the BQ data
    - A vast amount of variables were derived from original responses to the BQ. These variables relate to a set of broad groups, namely the respondent's background, education/training, earnings and skill use.
        - The majority of these variables were computed automatically during each pass over the data. These were based on definitions provided by the OECD and other partners of the Consortium.
        - Derived earnings variables were directly derived in the case of detailed responses or imputed from broad categories, and merged to the files.

        - Skill use derived variables were based on IRT estimation procedures, computed, and merged to the files.
    - A set of coarsened variables was scripted at the IEA DPC to cater for countries' needs to protect the confidentiality of respondents' information in the database. For these variables (suffix "_C"), one of three types of coarsening was applied: i) top coding, ii) categorization, or iii) collapsing of existing categories into a smaller set.
    - Finally, a set of "trend" variables was derived by ETS and provided to the IEA DPC as mergeable files (suffix "_T"). These trend variables relate to variables collected in the same or similar way as the ALL and IALS surveys; PIAAC variables were recoded to match the metric or coding schemes used in ALL and IALS in order to be comparable across surveys.
    - A small number of the derived BQ, trend and coarsened variables were computed under the responsibility of countries because the Consortium was not given access to the full source information required for the derivation. These variables were provided as mergeable files, validated and merged at the IEA DPC.

- Derivation of variables from the actual responses to the reading components items
    - At the time of data collection, three different types of response value schemes were used on the response capture sheets for print vocabulary, sentence processing and passage comprehension. During the data processing a response key was programmatically applied and used to assign actual responses (variables ending in "A") to scored responses (ending in "S") for all reading component items by mapping the correct distractor to code 1 = "correct" and other distractors to 7 = "incorrect."

- Derivation of variables from problem-solving unit responses
    - The PIAAC CBA system stored rich auxiliary information that provided indicators of respondents' actions during the cognitive assessment. At the time of collection, a large number of aggregate variables and interim scores were exported and processed. Following the data collection, "total correct scores" were derived and integrated into the master databases.

- Derivation of scale scores
    - PIAAC cognitive item responses were calibrated, analyzed and scaled. This process resulted in a set of 10 plausible values for each domain (literacy, numeracy and problem solving) plus one additional variable indicating the availability of plausible values for a particular respondent given the design and path.

13.4.5 Production of the preliminary national and international databases

The data finalization phase transitioned data from the internal IEA DPC processing systems to data products ready for further use by the Consortium, the OECD or the participating countries. The final processing phase further repeated many of the checks implemented at earlier stages to ensure that automated or incidental data editing did not introduce any inconsistencies, for example out-of-range codes, into the data. In addition, a set of additional checks was conducted that ensured data integrity after all cleaning steps had been run through and before export to the different final formats took place.

For example, checks ensured that the variable widths and types in the codebooks were defined wide enough to actually hold the data in the national master database.

At this stage, a single international codebook was used to describe and document the data. Widening conversions were applied consistently across all countries in case one or more countries extended the width of a variable in their national database's codebook (e.g., with respect to currency values). The final international master database held 1,712 international variables for each participating country. Codebook information for nationally adapted or extended variables was taken from the national databases originally submitted by countries.

In all, the 33 datasets present in the IEA DME software and database at the time of data capture were processed and eventually resulted in the following six export file types, each produced in both SPSS as well as SAS format:

- *PRGxxxMS.sav/.sas7bdat*: The main analysis file with all originally collected and derived variables, international as well as national.

- *PRRxxxMS.sav/.sas7bdat*: An auxiliary file holding reliability scores for the core and literacy/numeracy booklets as well as responses captured for reading components. The PRR file includes a true subset of the variables in PRG but with values from the reliability scoring process.

- *PAGxxxMS.sav/.sas7bdat*: A flat file with scores from the cross-country reliability study, main scoring.

- *PARxxxMS.sav/.sas7bdat*: A flat file with scores from the cross-country reliability study, reliability scoring.

- *PSDxxxMS.sav/.sas7bdat*: A flat file encompassing sample design variables. This file included a true subset of variables as well as all records from the PRG file and was mainly used by Westat or countries in the process of weighting.

- *PRLxxxMS.sav/.sas7bdat*: A flat file for the CAPI event log.

Data files were exported separately by country. This allowed for the provision of files to the Consortium as well as to individual countries on a rolling basis. The placeholder "xxx" used in the file names above corresponds to operational identifiers based on ISO 3166.

SPSS data files were standard, Windows-based *.sav* files and encoded in Unicode (UTF-8). SPSS data files included full dictionary information from the applicable metadata maintained in the codebooks including variable types and formats, variable labels, value labels (including any labels for missing values), missing value definitions and variable measurement levels. SAS-formatted files were standard, compressed *.sas7bdat* data files for Windows environments and encoded in Unicode (UTF-8). Variable types, widths, decimals and labels were assigned to all variables according to the labels defined in the metadata. SAS does not provide for a way to permanently store value labels on the file. Therefore, each file in SAS format was accompanied by an equivalently named *.sas* file which could be used to assign formats (value labels) to working files. Missing values represented as SAS missing values were programmatically mapped to either numerical missing values in the case of SPSS or a reduced set of special missing values in the case of SAS.

To allow for the export of data products for the various data users and stakeholder, data files could be produced according to three export profiles:

- Profile 1 for international analysis, weighting and archiving
    - This export profile retained all international and national variables originally submitted or derived on the data file.
    - These full information files were made available only to the Consortium partners who required access to the data as well as the OECD. These files were kept strictly confidential and were not shared beyond the group of organizations and individuals involved in the analysis and weighting.
    - This profile included all records originally submitted by a country.
- Profile 2 for the release of national databases to countries
    - This export profile maintained the vast majority of international and national variables. It excluded a small set of internal, interim or redundant variables produced as part of the scaling and analysis process and only relevant for the purpose or archiving.
    - This profile was provided only to the concerned countries.
    - This profile included all records originally submitted by a country.
- Profile 3 for public use
    - This export profile, by default, maintained all international variables approved for release by the BPC as part of the public-use file.
    - Any and all national variables were dropped.
    - For this profile, all international variables earmarked for suppression by a country were blanked (i.e., set to the appropriate missing value for all cases).
    - This profile only included records with the PIAAC in-sample flag (INPIAAC) equal to 1.

Each data exported was uniquely identified by an export data and an export version variable in the data files. These two variables allowed analysts to compare the data version underpinning the current work. In terms of data flow, the IEA DPC, as a subcontractor, provided all data products exclusively to ETS followed by quality control there. Subsequent data releases to other Consortium partners, the OECD, and participating countries were managed by ETS. An alternative data exchange protocol was used in the case of Australia to account for special regulations pertaining data security.

13.5 Data review and finalization

Following the initial data cleaning process described above, an iterative process of data review and correction began within the Consortium and later involved the participating countries as well as the OECD Secretariat. Integrating, verifying and, where necessary, updating the above stated groups of variables as well as the implementation of countries' feedback on their national databases all occurred under a tight timeline and included multiple data sendouts and review rounds. The general principle followed was that data collected, cleaned or derived by one party (e.g., the participating country or a Consortium partner) was reviewed by at least one other partner as well as the concerned country. Building and verifying the national and international databases was a collaborative process involving the specific expertise, knowledge and experience of the surveys designers, stakeholders and national project teams.

The list below presents the key data products and times in the process of reviewing and finalizing national and international databases for the majority of countries.

- Preliminary international database (July 2012)
    - The IEA DPC provided a preliminary international database including data from 20 countries to the Consortium for internal review and to ensure that all processes and procedures for analyzing Main Study data were in place.
    - This database included originally submitted, initially cleaned, and where applicable, perturbed data. Further, this database contained the design weights provided by countries.
    - A series of country-by-country updates to the preliminary international database and initial versions for two late-submitting countries were issued between July and November 2012 in parallel to data cleaning and initial weighting efforts.
- First international database (December 2012)
    - The IEA DPC provided a first international database including 22 countries’ data for analysis to the Consortium.
    - This database included weights, replicate weights, and a basic set of scripted derived variables.
- Second international database (January 2013)
    - The Consortium completed the initial data analysis and generated the majority of derived variables and plausible values for 22 countries.
    - This database was shared with the OECD in order to prepare international reporting.
- Release of preliminary national databases (January 2013)
    - At the end of January 2013, the Consortium released cleaned, weighted and analyzed national data to countries for review and approval. The microdata files were accompanied by summary data tables.
- Review of preliminary national databases (February to June 2013)
    - This period, originally scheduled until the end of May 2013, was intended for countries to review records and variables included in their cleaned, weighted and analyzed national databases.

- As a result of countries' review of their respective national database, the Consortium's own observations, and the initial reporting work at the OECD, the Consortium and the OECD agreed on data changes and error corrections to be applied commonly for all or just individual countries in order to improve the validity and quality of the data. Such changes related to:

    - repeated or corrected coding of occupational information with initially insufficient reliability or agreement with external data sources (e.g., labor force surveys);
    - minor corrections to the mapping of national educational attainment variables to international ISCED levels in light of discrepancies with other data collections (such as the OECD's Education at a Glance) in some countries;
    - correction of outliers in earning variables for some countries;
    - assignment of valid skip codes for skill use, earnings, reading components outcome variables, and a few other variables given that the original variables were not applicable to the entire survey population;
    - corrections to country-specific or general scripted derived variables; and
    - numerous label changes to better describe and reflect the content and scope of variables.

- The correction of data in some cases required the reanalysis of the cognitive data; resulting updates to scale scores and other measures were reflected in the concerned national databases.

- Countries were further asked to identify variables for suppressions and coarsening in any public-use data file releases on the basis of a preliminary list of variables earmarked for inclusion in such files. By mid-February, countries provided the Consortium with lists of variables coarsened. From this, the OECD selected a set of coarsened variables to be included for all countries. By the end of March, countries provided the Consortium with a list of variables to be suppressed from the now complete set of variables intended for the public-use data.

• Release of restricted international database through the Data Explorer (April 2013)

    - Following the initial batch of corrections and updates, the IEA DPC finalized a third international database including 22 countries.
    - The third international database was shared with the OECD in order to continue the preparation of international reporting.
    - The database was exposed to participating countries via an initial, secure version of the PIAAC Data Explorer at the time of a training workshop delivered to NPMs. Access to this version of the Data Explorer was restricted to countries, the OECD and the Consortium partners. Countries were identified by codes rather than clear text names.

• Release of draft national public-use files (June 2013)

  - Following the earlier corrections, the IEA DPC produced a draft of the public-use file for each country that reflected the respective national suppressions.

  - Countries were asked to verify the contents and accurate suppression.

- Finalization of the international database (June/July 2013)

  - The international database for 22 countries was finalized at the IEA DPC. The Consortium applied last-minute tweaks to variable and value labels, as well as to missing value schemes.

  - This database was shared with the OECD in order to produce an updated draft of the international report.

- Release of unrestricted Data Explorer and public-use data to all countries (July 2013)

  - Countries received unrestricted but embargoed access to the PIAAC Data Explorer, that is, with country names unmasked, and the public-use data files for all other countries in order to advance and finalize work on national reports.

  - In August, data for France was additionally released to countries.

- Release of the international report and a public-use international database (October 2013).

  - The public-use version of the international database was scheduled to be released in parallel to the initial international report for PIAAC.

Chapter 14: Sampling Design

Leyla Mohadjer, Tom Krenzke and Wendy Van de Kerckhove, Westat

This chapter presents information about the PIAAC Main Study sample design and selection results. Participating countries were required to develop their sample design and selection plans according to the standards provided in the PIAAC Technical Standards and Guidelines (TSG) and to submit their plans to the Consortium for approval. The sample design plans included information about sampling frames and their coverage, providing descriptions of the national sample designs that included stages of sampling, probabilities of selection, sampling units and sample sizes. The sample selection plans included detailed information about the processes for sample selection at each stage of sampling. In addition, the countries were required to complete and submit quality control sample selection forms to the Consortium to verify that the sample selection was conducted in an unbiased and randomized way consistent with PIAAC standards.

The target population for PIAAC consists of all noninstitutionalized adults between age 16 and 65 (inclusive) who reside in the country (meaning their usual place of residency is in the country) at the time of data collection. Countries were allowed to expand the target population to include additional subpopulations of interest to the country as long as they followed the TSG on such supplementation. Section 14.1 provides more detail on the PIAAC target population and the national target populations if expanded beyond the PIAAC standard definition. Section 14.2 contains information about the sources of country sampling frames and their coverage of the target population.

The TSG allowed each country to choose a sample design and selection approach that is most optimal and cost effective as long as the design applies full selection probability methods to select a representative sample from the PIAAC target population. Descriptions of the standard PIAAC and national sample designs and probabilities of selection are given in section 14.3. The definition of sampling units and sample selection methods are provided in section 14.4. Section 14.5 contains the PIAAC target sample sizes and describes the process applied to determine the initial sample sizes. Sample selection results and a summary of the sampling quality control procedures are given in section 14.6 and section 14.7, respectively. Finally, section 14.8 provides a brief description of the incentive plans for PIAAC.

14.1 Target population and sampling frame

A clear and precise definition of the target population is necessary to ensure that the population of interest is adequately covered by each participating country and to maintain consistency and comparability across countries. The PIAAC target population consists of all noninstitutionalized adults between age 16 and 65 (inclusive) who reside in the country (usual place of residency is in the country) at the time of data collection. Adults were to be included regardless of citizenship, nationality or language (standard 4.1.1). The target population excludes adults in institutional collective dwelling units (or group quarters) such as prisons, hospitals and nursing homes, as well as adults residing in military barracks and military bases. However, full-time and part-time members of the military who do not reside in military barracks or military bases are included in the target population.

Adults in other noninstitutional collective dwelling units (or group quarters), such as workers' quarters or halfway homes, are also included in the target population. This includes adults living at school in student group quarters such as a dormitory, fraternity or sorority. Adults who were unable to complete the assessment because of a hearing impairment, blindness/visual impairment or physical disability are considered in scope; however, they were excluded from PIAAC response rate calculations because the assessment does not accommodate such situations.

The target population does not cover the entire geography area for the following countries:

- Belgium – The target population consists of Flanders, which is in the northern portion of the country.
- Cyprus[1] – The target population consists of the area under the effective control of the Government of the Republic of Cyprus, which includes the districts of Nicosia (part), Limassol, Larnaca (part), Paphos and Famagusta (part).
- Russian Federation[2] – The target population does not include Moscow or Moscow Region.

Some countries expanded the target population to include additional subpopulations of interest to the country. These country-specific supplemental samples, approved by the Consortium, are presented in Table 14-1 below.

Table 14-1: Country-specific samples

Country	Specific samples
Australia	Persons aged 15 and 66-74
Denmark	PISA 2000 survey respondents aged 26-27

Some countries elected to oversample portions of the target population. The oversamples approved by the Consortium are presented in Table 14-2 below.

[1] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[2] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 14-2: Countries conducting oversampling

Country	Group oversampled
Australia	Persons living in certain states/territories
Canada	Individuals aged between 16 and 24 inclusive in British Columbia; Linguistic minorities (English in Québec, French elsewhere) in New Brunswick, Québec, Ontario and Manitoba; Métis in Ontario; Aboriginal individuals in Québec, Ontario, Manitoba, Saskatchewan, Alberta, British Columbia and Yukon Territory; and Recent immigrants (living in Canada since 2002 or after) in Québec, Ontario, Alberta and British Columbia
Czech Republic	Persons aged 16-29
Denmark	Persons aged 55-65 years and immigrants 16-65
Germany	Persons aged 26-55 living in former East Germany or former East Berlin [1]
Poland	Persons aged 19-26

[1] For national purposes; not included in the international data.

14.2 Sampling frames and their coverage

The sampling frame is the list from which the sample is selected, so the quality of the sampling frame affects the quality of the sample. In addition, adequate information on the frame must be available to conduct sampling, data collection, weighting and nonresponse bias analyses. Most countries with multiple stages of selection had specified multiple frames. Those frames were reviewed by the Consortium to ensure they included sufficiently reliable information for sampling individual units and ultimately locating individuals for the interview and assessment. Section 14.2.1 provides information about the sampling frames used at each stage of selection, while section 14.2.2 contains information about the coverage of these frames.

In PIAAC, the noncoverage rate, combined over all stages of sampling, could not exceed 5% (standard 4.1.2). Thus the sampling frames for each country were required to include 95% or more of the standard PIAAC target population. Frame noncoverage rates (see section 14.2.2) were limited as much as possible so that no extensive biases are introduced as a result of noncoverage of some subgroups of the population.

14.2.1 Sampling frames

PIAAC standards require that sampling frames be up to date and include only one record for each member of the target population. Countries had to examine their sampling frames and eliminate duplicate records when lists were combined to create a sampling frame. Countries were required to assess the extent of duplication and the proportion of out-of-scope units on the frame and, if necessary, develop a plan to correct these problems. In addition, countries also evaluated and developed plans to address any noncoverage in the frame that was not addressed in the documentation of country-specific exclusions (see Table 14-6). The methodology used to create these frames was also reviewed by the Consortium.

Multistage sample designs required a sampling frame for each stage of selection. Some countries used national population registries as sampling frames, which contain useful variables for stratification, weighting and nonresponse bias analyses. If the country had a list of residents that was of sufficient quality, no frame of households or household sampling was necessary.

However, some countries' lists of residents used for the study did not completely cover the PIAAC target population (e.g., the lists may have excluded nonnationals/noncitizens), complicating their use as a sampling frame. See Table 14-3 for the full list of sampling frames employed by countries with population registry samples.

Table 14-3: Sampling frames for countries with population registry samples

Country	Sampling frame		
	Stage 1	**Stage 2**	**Stage 3**
Austria	Population registry, 2011		
Denmark	Population registry, 2011		
Estonia	Population registry, 2011		
Finland	Statistics Finland's population database (based on the Central Population Register), 2011		
Flanders (Belgium)	Population registry, 2011		
Germany	German Census Bureau frame of communities, 2011	Local population registries, 2011	
Italy	National Statistical Institute of Italy frame of municipalities, 2011	Household registries held by municipalities, 2011	Population registries, 2011; combined with field enumeration
Japan	Resident registry, 2011	Resident registry, 2011	
Netherlands	Population registry, 2011		
Norway	Population registry, 2011		
Poland	Population registry, 2011	Population registry, 2011	
Slovak Republic	Population registry, 2011	Population registry, 2011	
Spain	Population registry, 2011	Population registry, 2011	
Sweden	Population registry, 2011		

indicates there is no such stage in the country's sample design.

Some countries have access to master samples used for national surveys. For example, Australia has a master sample of dwelling units (DUs) already in use by governmental surveys that was also used for PIAAC. Similarly, Australia and France have master samples of area primary sampling units (PSUs). See Table 14-4 for more information on how master samples were employed by participating countries.

Table 14-4: Sampling frames for countries using master samples

Country	Sampling frame			
	Stage 1	Stage 2	Stage 3	Stage 4
Australia	Bureau of Statistics population survey master sample, 2006	Bureau of Statistics population survey master sample, 2006	Bureau of Statistics population survey master sample, 2006	Field enumeration
France	Master sample from census data file, 1999	Individual taxation file, 2011		

For multistage area sample designs in which a registry is not being used, listing procedures are necessary to create a frame of households within the selected geographic clusters. A frame of geographic clusters can be formed by combining adjacent geographic areas, respecting their population sizes and taking into consideration travel distances for interviewers. Table 14-5 contains sampling frames for the remaining countries without registries using area sample designs for PIAAC. The exception is that Cyprus[3] is included in Table 14-5 among the countries without population registries, even though it did not use an area sample design, Cyprus did not require listing procedures because its sample frame for the first stage was a list of households from the Statistical Service Census 2001, updated with information from the 2010 Electricity Authority Household Registry.

[3] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 14-5: Sampling frames for countries without population registries and master samples

Country	Sampling frame			
	Stage 1	Stage 2	Stage 3	Stage 4
Canada	Short-form Census returns and National Household Survey returns for some oversamples, 2011	Short-form Census returns and National Household Survey returns for some oversamples, 2011	Field enumeration	
Cyprus [4]	List of households from the Statistical Service Census 2001, updated with information from the 2010 Electricity Authority Household Registry	Field enumeration		
Czech Republic	Territorial Identification Register of Buildings and addresses (UIR-ADR), 2010	Territorial Identification Register of Buildings and addresses (UIR-ADR), 2010	Field enumeration	Field enumeration
England (UK)	Royal Mail list of UK Postal Sectors, 2011	Royal Mail PAF residential file, 2011	Field enumeration	Field enumeration
Ireland	Small Area classifications, 2006	2011 Census	Field enumeration	
Korea	2010 Census	2010 Census	Field enumeration	
Northern Ireland (UK)	NI(POINTER) database, 2011	Field enumeration	Field enumeration	
Russian Federation[5]	Federal State Statistics Service, data of the national survey organizations, 2010	Federal State Statistics Service, data of the national survey organizations, 2010	Official data of urban districts, 2010	Field enumeration
United States	Census Bureau Population Estimates, 2008	2000 Census Bureau Summary File 1 (SF1), 2000; updated with data from the United States Postal Service 2010	Field enumeration	Field enumeration

indicates there is no such stage in the country's sample design.

14.2.2 Noncoverage of the target population

As mentioned earlier, the noncoverage rate for PIAAC, combined over all stages of sampling, may not exceed 5% (standard 4.1.2), and thus the sampling frames for each country were required to include 95% or more of the standard PIAAC target population. All exclusions to the core PIAAC target population, whether or not they exceed the threshold, were reviewed by the Consortium. Exclusions are acceptable only if they occur because of operational or resource

[4] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[5] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

considerations such as excluding persons in hard-to-reach areas. The Consortium asked that each country identify to the extent possible exclusions before sample selection. Adjustments for any noncoverage of the target population in each country was made through benchmarking during the weighting process (see Chapter 15). A complete list of exclusions for countries using population registries is presented in Table 14-6; Table 14-7 includes a similar list for countries not using population registries.

In addition to PIAAC eligible persons not included in sampling frames, persons that were included in the frame but in practice were impossible to be interviewed were treated as exclusions conditional on the total exclusion rate staying at or below 5%. Chapter 16 provides more information about this group, with Table 16-2 showing the overall exclusion rate for each country.

Table 14-6: Portion of target population not covered by Main Study sampling frames for countries using population registries

Country	Percentage of target population not covered*	Group not covered
Austria	0.6%	Undocumented immigrants
Denmark	< 0.1%	Undocumented immigrants
Estonia	2.8%+	Persons without a detailed address; undocumented immigrants (no estimate provided)
Finland	0.2%	Undocumented immigrants; asylum seekers
Flanders (Belgium)	1.0%	Undocumented immigrants
Germany	0.5%	Undocumented immigrants
Italy	0.8%+	Adults in noninstitutional group quarters; undocumented immigrants (no estimate provided)
Japan	2.2%	Nonnationals; undocumented immigrants
Netherlands	0.9%	Undocumented immigrants
Norway	0.4%	Undocumented immigrants
Poland	0.8%	Foreigners staying in Poland fewer than 3 months; nonregistered immigrants
Slovak Republic	0.1%	Undocumented immigrants
Spain	0.0%	None
Sweden	< 1.0%	Undocumented immigrants

* The noncoverage rate accounts for excluded subpopulations such as undocumented immigrants or noninstitutionalized collective DUs, with the exception that the homeless are not being considered part of this rate. Other exclusions that will occur as a natural part of the survey process are not included in the expected noncoverage rate.

Table 14-7: Portion of target population not covered by Main Study sampling frames for countries not using population registries

Country	Percentage of target population not covered*	Group not covered
Australia	3.3%	Persons living in very remote areas, discrete indigenous communities (DIC), or noninstitutional special dwellings; non-Australian diplomats, their staff and household members of such; members (and their dependents) of non-Australian defense forces
Canada	1.8%	Residents of smallest communities in the northern territories; residents of remote and very low population density areas in provinces; and persons living in noninstitutional collective dwellings, other than students in residences.
Cyprus [6]	< 2.0%	Persons living in houses built after December 2010
Czech Republic	1.8%	Professional armed forces; municipalities with < 200 habitants
England/Northern Ireland (UK)	2.0%	Individuals living in private residences that are not listed on the "residential" version of the Postal Address File (PAF) or, in Northern Ireland (UK), not listed on the NI(POINTER) database
France	< 2.6%	Young adults who have never claimed any income and are not attached to their parents households; undocumented immigrants
Ireland	0.4%	Some mobile dwellings
Korea	2.4%	Small islands residents
Russian Federation[7]	1.5%	Chechnya region
United States	0.1%	People in large gated communities

* The noncoverage rate accounts for excluded subpopulations such as undocumented immigrants or noninstitutionalized collective DUs, with the exception that the homeless are not being considered part of this rate. Other exclusions that will occur as a natural part of the survey process are not included in the expected noncoverage rate.

14.3 National sample designs

The PIAAC standard sample design is a self-weighting design of persons (or of households, for countries without person registries). A self-weighting design is achieved when each sample person (or household, if sampling dwelling units) has an equal probability of selection (standard 4.4.3). For countries that are geographically large, the typical sample design is a stratified multistage clustered area sample. For participating countries that are geographically small, the sample design had less clustering and fewer stages of sampling. Also, several countries had lists of households or persons already available from national registries or registries managed by municipalities.

[6] Please refer to notes A and B regarding Cyprus in the Note to Readers section of this report.

7 Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

The TSG allow each country to choose a sample design and selection approach that is most optimal and cost effective as long as the sample design applies full selection probability methods. Each participating country was required to produce a probability-based sample, representative of the target population of the country. The PIAAC standards require probability-based samples because they are essential for two main reasons. First, probability sampling encompasses a set of designs that leads to a variety of unbiased sampling approaches that allow analysts to generalize the results to the target population. Second, measures of precision related to survey estimates (i.e., standard errors, margins of error, confidence intervals) can be computed under a probability design only. Hence, statistical tests for differences between survey estimates are possible only under a probability-based design.

The PIAAC standard probabilities of selection as applied to each country's design are presented in section 14.3.1. Section 14.4.1 presents the sample units selected at each stage of selection, while section 14.4.2 presents the sample selection methods. The factors contributing to the sample size determination in each country, and the sample sizes, are presented in section 14.5.

14.3.1 Probabilities of selection based on PIAAC standard design

Each person in the PIAAC target population must have a nonzero probability of selection resulting from the application of established and professionally recognized principles of scientific sampling (standard 4.4.1). As the ultimate sampling unit, each person in the PIAAC target population must have a calculable nonzero probability of selection. That is, every in-scope person must have a chance of being selected into the PIAAC sample. The following presents the PIAAC approach that was recommended for selecting the ultimate sampling unit for one-, two-, three-, and four-stage sample designs, respectively. The approach is based on PIAAC standards and guidelines. Countries were sent the formulas prior to their sample selection process, and they were asked to confirm or to provide formulas showing their deviations from the self-weighting design. The Consortium conducted checks during and after sample selection. Some countries deviated from these formulae due to oversampling (as given in Table 14-2) or alternative sampling formulas. Table 16-8 provides the variation of the base weights, which identifies the countries that achieved self-weighting or near self-weighting designs (a coefficient of variation of less than 0.05). Among the 14 registry countries, self-weighting or near self-weighting designs were achieved by Austria, Flanders (Belgium), Estonia, Finland, Japan, Netherlands, Norway, Slovak Republic and Sweden. Among the nine screener countries (treating England and Northern Ireland as separate designs), self-weighting or near self-weighting of dwelling units was achieved by Cyprus[8] and the United States.

One-stage sample designs

For a one-stage sample design without any explicit stratification, let

n = total number of persons to be sample, and

N = total number of eligible persons.

The probability of selecting person l is $r = n/N$.

[8] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Austria was the only country that adapted a one-stage sample design with no explicit stratification.

For a one-stage stratified sample design, let

n_h = number of persons to be sampled in stratum h; and

N_h = number of eligible persons in stratum h.

Further, let $r = n/N$, then the probability of selecting person l in strata h is

$P_{hl} = r$.

The sample size is allocated to strata as

$n_h = P_{hl} \times N_h = r \times N_h$.

Seven countries used a one-stage stratified sample design: Flanders (Belgium), Denmark, Estonia, Finland, Netherlands, Norway and Sweden.

Two-stage stratified probability proportionate to size designs

The formulae for the standard PIAAC selection probabilities for each stage are given below.

For the first-stage sample of primary sampling units (PSUs) in the remaining countries, let

m_h = number of PSUs to be sampled in stratum h;

MOS_{hi} = measure of size for PSU i in stratum h; and

I_{psu}^{h} = sampling interval for the selection of PSUs in stratum h.

The probability of selecting PSU i in stratum h is

$$P_{hi} = \frac{m_h \times MOS_{hi}}{\sum_{i \in h} MOS_{hi}} = \frac{MOS_{hi}}{I_{psu}^{h}}$$

For the second-stage sample of persons, let

n = total number of persons to be sampled;

N = total number of eligible persons;

n_{hi} = number of persons to be sampled in PSU i of stratum h; and

N_{hi} = number of eligible persons in PSU i of stratum h.

Let $r = n/N$, then the *conditional* probability of selecting person l in PSU i of stratum h is

$$CP_{hil} = \frac{r}{P_{hi}} = r \times \frac{I_{psu}^{h}}{MOS_{hi}}$$

The *overall* probability of selecting person l in PSU i of stratum h is

$P_{hil} = P_{hi} \times CP_{hil} = r.$

The sample size in PSU i of stratum h is

$$n_{hi} = CP_{hil} \times N_{hi} = r \times \frac{\sum_{i \in h} MOS_{hi}}{m_h} \times \frac{N_{hi}}{MOS_{hi}} = r \times I_{psu}^{h} \times \frac{N_{hi}}{MOS_{hi}}$$

Seven countries used a two-stage stratified sample design: Cyprus,[9] France, Germany, Japan, Poland, Slovak Republic and Spain. Poland's weights varied due to oversampling and by applying an alternative design implementation strategy. France used a different approach that followed balance sampling (Deville & Tillé, 2004 and Tillé, 2006) that resulted in varying base weights. Germany's design included deep stratification in the context of Cox (1987) and included simulated values for probabilities of selection due to a sampling-related problem. Spain's weights varied due to applying an alternative design implementation strategy.

Three-stage stratified probability proportionate to size (PPS) designs

In a three-stage stratified PPS design, PSUs are selected with a probability proportionate to a measure of size as described below.

For PSU selection in the training countries, let

m_h = number of PSUs to be sampled in stratum h;

MOS_{hi} = measure of size for PSU i in stratum h; and

I_{psu}^{h} = sampling interval for the selection of PSUs in stratum h.

The probability of selecting PSU i in stratum h is

$$P_{hi} = \frac{m_h \times MOS_{hi}}{\sum_{i \in h} MOS_{hi}} = \frac{MOS_{hi}}{I_{psu}^{h}}$$

For the second stage sample of dwelling units (DUs), let

d = total number of housing units to be sampled;

D = total number of housing units in the sampling frame;

d_{hi} = number of housing units to be sampled in PSU i of stratum h; and

D_{hi} = number of housing units in PSU i of stratum h.

Let $r = d/D$, then the *conditional* probability of selecting housing unit k from PSU i in stratum h is

[9] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

$$CP_{hik} = \frac{r}{P_{hi}} = r \times \frac{I_{psu}^{h}}{MOS_{hi}}$$

The *overall* probability of selecting housing unit k in PSU i of stratum h is

$$P_{hik} = P_{hi} \times CP_{hik}$$

The DU sample size in a PSU is

$$d_{hi} = CP_{hik} \times D_{hi} = r \times \frac{\sum_{i \in h} MOS_{hi}}{m_h} \times \frac{D_{hi}}{MOS_{hi}} = r \times I_{psu}^{h} \times \frac{D_{hi}}{MOS_{hi}}$$

For person selection, let

n_{hik} = number of persons to be sampled from housing unit k in PSU i of stratum h; and

N_{hik} = total number of eligible persons in housing unit k of PSU i in stratum h.

The *conditional* probability of selecting person l from housing unit k in PSU i of stratum h is

$$CP_{hikl} = \frac{n_{hik}}{N_{hik}}$$

The *overall* probability of selecting person l in housing unit k of PSU i of stratum h is

$$P_{hikl} = P_{hi} \times CP_{hik} \times CP_{hikl} = r \times \frac{n_{hik}}{N_{hik}}$$

Canada, Ireland, Italy, Korea and the Northern Ireland design stratum of the United Kingdom all used a three-stage stratified PPS design. Canada's weights varied due to oversampling. Ireland implemented a sample size-based design in lieu of rate-based design, which caused some variation in the base weights. Italy, Korea and Northern Ireland (UK) each applied an alternative design implementation strategy that caused variation, excessive in the case of Northern Ireland (UK), in the resulting base weights.

Four-stage stratified probability proportionate to size designs

Within the four-stage stratified PPS sample design, PSUs and secondary selection units (SSUs) are selected with a probability proportionate to a measure of size (MOS) as described below.

For PSU selection in the reaming countries, let

m_h = number of PSUs to be sampled in stratum h; and

MOS_{hi} = measure of size for PSU i in stratum h.

The probability of selecting PSU i in stratum h is

$$P_{hi} = \frac{m_h \times MOS_{hi}}{\sum_{i \in h} MOS_{hi}}$$

For SSU selection, let

q = total number of SSUs to be sampled;

MOS_{hij} = measure of size for SSU j of PSU i in stratum h; and

I_{SSU} = sampling interval for the selection of SSUs.

The *conditional* probability of selecting SSU j from PSU i in stratum h is

$$CP_{hij} = \frac{q \times \left(\frac{MOS_{hij}}{P_{hi}}\right)}{\sum_{hij}\left(\frac{MOS_{hij}}{P_{hi}}\right)} = \frac{MOS_{hij}/P_{hi}}{I_{SSU}}$$

For DU selection, let

d = total number of housing units to be sampled;

D = total number of housing units in the sampling frame;

d_{hij} = number of housing units to be sampled in SSU j of PSU i of stratum h; and

D_{hij} = number of housing units in SSU j of PSU i of stratum h.

Let $= d/D$, then the *conditional* probability of selecting housing unit k from SSU j of PSU i in stratum h is

$$CP_{hijk} = \frac{r}{P_{hi} \times CP_{hij}} = \frac{r \times I_{SSU}}{MOS_{hij}}$$

The *overall* probability of selecting housing unit k in SSU j of PSU i of stratum h is

$$P_{hijk} = P_{hi} \times CP_{hij} \times CP_{hijk} = r$$

The DU sample size in a SSU is

$$d_{hij} = CP_{hijk} \times D_{hij} = r \times I_{SSU} \times \frac{D_{hij}}{MOS_{hij}}$$

For person selection, let

n_{hijk} = number of persons to be sampled from housing unit k of SSU j in PSU i within stratum h; and

N_{hijk} = total number of eligible persons in housing unit k of SSU j in PSU i within stratum h.

The *conditional* probability of selecting person l from housing unit k of SSU j in PSU i within stratum h is

$$CP_{hijkl} = \frac{n_{hijk}}{N_{hijk}}$$

The *overall* probability of selecting person l from housing unit k of SSU j in PSU i within stratum h is

$$P_{hijkl} = P_{hi} \times CP_{hij} \times CP_{hijk} \times CP_{hijkl} = r \times \frac{n_{hijk}}{N_{hijk}}$$

Australia, the Czech Republic, the Russian Federation,[10] the England design stratum of the United Kingdom, and the United States used a four-stage stratified PPS sample design. The Czech Republic conducted oversampling and also implemented a sequential selection design strategy that caused excessive variation in the resulting base weights. England (UK) had variation in its base weights due to implementing a selection process that is different from the one outlined with the above formulae.

14.4 Sample units and sample selection methods

14.4.1 Sample units

Because Austria, Flanders (Belgium), Denmark, Estonia, Finland, Netherlands, Norway and Sweden all implemented a one-stage sample design, they have only one sample unit: persons. The sampling units for countries with two-, three-, and four-stage sample designs are shown in Tables 14-8 to 14-10, respectively.

Table 14-8: Main study sample units for countries with two stages of sampling

Country		Stage 1	Stage 2
Cyprus [11]		Households	Persons
France		Area PSUs	Persons
Germany		Communities	Persons
Japan		Cho/Chome/Aza administrative districts	Persons
Poland	Urban	Towns/Cities	Persons
	Rural	Towns/Villages	Persons
Slovak Republic		Municipalities	Persons
Spain		Area PSUs	Persons

Note: “Area PSUs” indicates primary sampling unit covers a geographic area not defined by a generic geographic terminology (towns, villages, etc).

[10] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[11] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 14-9: Main study sample units for countries with three stages of sampling

Country	Stage 1	Stage 2	Stage 3
Canada	Area PSUs	DUs	Persons
Ireland	Area PSUs	Households	Persons
Italy	Municipalities	Households	Persons
Korea	Enumeration districts	DUs	Persons

Note: "Area PSUs" indicates primary unit covers a geographic area not defined by a generic geographic terminology (towns, villages, etc).

"DUs" indicates dwelling units; "Households" are occupied DUs.

Table 14-10: Main Study sample units for countries with four stages of sampling

Country	Stage 1	Stage 2	Stage 3	Stage 4
Australia	Area PSUs	Blocks	DUs	Persons
Czech Republic	Districts (sub-regions)	Streets	DUs	Persons
England/Northern Ireland (UK)	Postal sectors Addresses	Addresses Households	Households Persons	Persons
Russian Federation[12]	Regions	Settlements	DUs	Persons
United States	Area PSUs	Area SSUs	DUs	Persons

Note: "Area PSUs" or "Area SSUs" indicates primary or secondary sampling unit covers a geographic area not defined by a generic geographic terminology (towns, villages, etc).

"DUs" indicates dwelling units; "Households" are occupied DUs.

14.4.2 Sample selection methods

Details regarding the selection methods for countries with one- or two-stage sample designs are presented in Tables 14-11 and 14-12, respectively.

Table 14-11: Main Study selection methods for countries with one stage of selection

Country	Description
Austria	Systematic random sample from a sorted list
Denmark	SRS within explicit strata
Estonia	Systematic random from a sorted list within explicit strata
Finland	Systematic random from a sorted list within explicit strata
Flanders (Belgium)	Systematic random from a sorted list within explicit strata
Netherlands	SRS within explicit strata
Norway	SRS within explicit strata
Sweden	SRS within explicit strata

Note: "SRS" indicates simple random sampling.

[12] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 14-12: Main Study selection methods for countries with two stages of selection

Country		Stage	Description
Cyprus [13]		1	Systematic random from a sorted list within explicit strata
		2	SRS of 1 person per household via pre-assigned selection grid
France		1	Systematic random from master sample IAAs (master sample selected using the balanced sampling algorithm, the "Cube" method, PPS (number of main residences in the IAA))
		2	Systematic random from a sorted list
Germany		1	Stratified, PPS (target population) with allocation by controlled rounding
		2	Two-phase sample. • Phase 1: The registries of the selected communities were asked to select an EPSEM sample of individuals. • Phase 2: Within each community, the individuals selected in Phase 1 were allocated to a matrix that was divided into six age groups x gender. Allocation of the Phase 2 sample size was done using an Iterative Proportional Fitting (IPF) procedure. The selection of persons within a community was done by systematic random sampling with a random start number and a sampling interval.
Japan		1	Systematic PPS (number of inhabitants age 15-64 as of March 2010) from a sorted list within explicit strata
		2	Systematic random from a sorted list
Poland	Urban	1	All towns/cities selected with certainty
		2	SRS within explicit strata
	Rural	1	PPS (population age 16-65) within explicit strata
		2	SRS without replacement of clusters of 8 persons in explicit strata
Slovak Republic		1	Systematic PPS (population age 16-65) from a sorted list within explicit strata
		2	Systematic random from a sorted list
Spain		1	Systematic PPS (population) from a sorted list within explicit strata
		2	Systematic random from a sorted list

Note: "SRS" indicates simple random sampling.

All countries with three- or four-stage designs selected samples of dwelling units before the enumeration and selection of persons within households. Although the goal was to select one person per household, the selection of more than one person per household was preferred for countries with a large variation in household size (standard 4.4.4). These include the Russian Federation[14] and the United States. Details regarding the selection methods for countries with three- or four -stage designs are presented in Tables 14-13 and 14-14, respectively.

[13] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[14] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 14-13: Main Study selection methods for countries with three stages of selection

Country	Stage	Description
Canada	1	Systematic PPS (2006 population counts) from a sorted list within explicit strata with Census Metropolitan Areas sampled with certainty
	2	Systematic random from a sorted list within explicit strata
	3	SRS of 1 person per household via pre-assigned hash number
Ireland	1	Stratified PPS (total dwellings)
	2	SRS
	3	SRS of 1 person per household
Italy	1	Systematic PPS (target population) from a sorted list within explicit strata
	2	Systematic random from a sorted list
	3	SRS of 1 person per household via selection grid is used if the household composition is different from the register; otherwise SRS from registry.
Korea	1	Systematic random sample from a sorted list within explicit strata
	2	Systematic random from a sorted list
	3	SRS of 1 person per household

Note: "SRS" indicates simple random sampling.

Table 14-14: Main Study selection methods for countries with four stages of selection

Country	Stage	Description
Australia	1	Systematic PPS (number of DU clusters) from a sorted list within explicit strata (subsample from master sample)
	2	Systematic PPS (number of DU clusters) from a sorted list (subsample from master sample)
	3	Systematic random from a sorted list
	4	SRS of 1 person per household
Czech Republic	1	Systematic PPS (number of inhabitants aged 16-65) from a sorted list within explicit strata
	2	Systematic PPS (number of address points)
	3	SRS; selected a "basic" sample of households to achieve the 5,000 completes plus an additional sample of households in which only 16- to 29-year-olds were sampled.
	4	SRS of 1 person per household
England (UK)	1	Systematic PPS (PAF single occupancy count) from a sorted list within explicit strata
	2	Systematic random from a sorted list
	3	SRS of 1 household at the sampled address using the Kish grid
	4	SRS of 1 person per household using the Kish grid
Northern Ireland (UK)	1	Systematic random from a sorted list
	2	SRS of 1 household at the sampled address using the Kish grid
	3	SRS of 1 person per household using the Kish grid
Russian Federation[15]	1	Systematic PPS (population in the region) from a sorted list within explicit strata
	2	Systematic PPS (target population) from a sorted list
	3	Systematic random from a sorted list
	4	SRS of 1 person for household sizes up to 4 (otherwise 2 persons) via pre-assigned selection grid
United States (USA)	1	Systematic PPS (population) within explicit strata
	2	Systematic PPS (number of DUs) from a sorted list
	3	Systematic random from a sorted list
	4	SRS of 1 person for household size up to 3 (otherwise 2 persons)

Note: "SRS" indicates simple random sampling.

[15] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Stratification combines sample units into homogeneous groups and reduces sampling variability between such groups and thus reduces the overall sampling variance associated with the resulting survey estimates. To maximize the benefit of stratification, stratification variables should be reliable and related to the survey outcome. Many of the countries utilizing population registries have the benefit of person-level characteristics available as stratification variables. The stratification and/or sorting variables for countries with one, two, three, and four stages of selection are detailed in Tables 14-15 to 4-18, respectively.

Table 14-15: Main Study stratification/sorting variables and methods for countries with one stage of selection

Country	Description
Austria	Sort by province, urban/rural, age, gender and citizenship
Denmark	Strata: age categories, immigration status
Estonia	Strata: gender and age categories Within strata: sort by region and age
Finland	Strata: native language (Finnish and other languages than Swedish, and Swedish) Within strata: sort by region, age, educational attainment, and gender
Flanders (Belgium)	Strata: province Within strata: sort by postal code, gender and age
Netherlands	Strata: municipality
Norway	Strata: level of education and age group
Sweden	Strata: gender, age, country of birth, level of education

Table 14-16: Main Study stratification/sorting variables and methods for countries with two stages of selection

Country		Stage	Description
Cyprus [16]		1	Strata: district, urban/rural classification Within strata: sort by geographic location
		2	None
France		1	Strata: administrative region (for master sample) Balancing variables: number of main residences, total income, number of DUs in rural, peri-urban, and urban areas.
		2	Stratified by housing (synthetic variable differentiating ordinary housing and communities) and sorted by department (administrative district).
Germany		1	Strata: region, urban/rural status (BIK) – approximately 1,000 strata cells
		2	None in Phase 1. In Phase 2, stratified by age group and gender, sorted by age.
Japan		1	Strata: region, urban/rural status; Sort by regional code
		2	Sort by address
Poland	Urban	1	Strata: size class
		2	Strata: age (19-26, other)
	Rural	1	Strata: region and size class
		2	Strata: age (19-26, other)
Slovak Republic		1	Strata: region, municipality size; Within strata: sort by number of age 16-65 in municipality
		2	Sort by gender and age
Spain		1	Strata: categories of municipality size Within strata: sort by population size
		2	Sort by gender and age

[16] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 14-17: Main Study stratification/sorting variables and methods for countries with three stages of selection

Country	Stage	Description
Canada	1	Stratify by province, urban/rural; sort by geographic order of PSUs and 2006 population counts
	2	Stratified by province/territory and urban/rural. Sort by geographic order (province/territory code, urban/rural, PSU ID, Census collection unit ID)
	3	None
Ireland	1	Strata: urban/rural status, and educational profile Within strata: sort by size of SAs
	2	None
	3	None
Italy	1	Strata: geographic regions of equal size Within strata: sort by the target population count of the PSUs
	2	None
	3	Random sort if selection from registry. If the household composition is different from the registry, persons are sorted by gender and age and the selection grid is used.
Korea	1	Strata: administrative districts Within strata: sort by enumeration district characteristics, such as townhouse versus apartment, percentage of 1-person household, education level, average age, percentage of people who are older than 60
	2	Sort by address
	3	None

Table 14-18: Main Study stratification/sorting variables and methods for countries with four stages of selection

Country	Stage	Description
Australia	1	Strata: state, part of state Within strata: serpentine sort by geography
	2	Serpentine sort by geography
	3	Serpentine sort by geography
	4	None
Czech Republic	1	Strata: region, municipality size Within strata: sort by code of location
	2	Sort by code of the street
	3	None
	4	Sort by year of birth
England (UK)	1	Strata: region, percentage living in social housing Within strata: sort by percentage of White British
	2	Sort by postcode and address number
	3	Sort by addresses (alphanumerically)
	4	Sort by first name
Northern Ireland (UK)	1	Sort by council ward, postcode within ward, and then alphanumerically within postcode
	2	Sort by addresses (alphanumerically)
	3	Sort by first name
Russian Federation[17]	1	Strata: macro regions Sort by federal county, population size for noncertainty PSUs
	2	Sort by type of settlement
	3	Sort by type of urban district (central/middle/outskirt)
	4	None
United States	1	Strata: region, metro area classification, race/ethnicity, income, percentage of the population that is foreign born
	2	Sort by geographic location
	3	Sort by geographic location
	4	None

14.5 Sample size determination

Adequate sample sizes are needed to establish stable item characteristics and to estimate separate population models for each tested language in a participating country. Population modeling is a critical step in obtaining appropriate proficiency values to be used in describing the distributions of skills in a country and in reporting national and subpopulation data.

The overall goal of the sample design for the Main Study was to obtain a nationally representative sample of the target population in each participating country that is proportional to the population across the country (i.e., a self-weighting sample design). As mentioned earlier, countries had the option of increasing sample sizes to obtain reliable estimates for groups of special interest (e.g., 16- to 29-year-olds), for geographic regions (e.g., states and provinces) or to extend the age range (e.g., 66-plus). However, the minimum sample size required was for a

[17] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

self-weighting design, and any sample size attributable to oversampling, or to subgroups outside of the PIAAC target population, was additional. PIAAC target sample sizes are presented in section 14.5.1.

To determine the initial sample size for the Main Study, the required number of assessments had to be adjusted to account for survey ineligibility and expected nonresponse to both the BQ/JRA and the assessment. For countries with a household screener, sample size goals had to be constructed for the screener to account for ineligibility and screener nonresponse, in addition to nonresponse to the BQ/JRA and assessment.

In most highly clustered surveys or those with a high degree of variability in sampling rates due to oversampling, initial sample sizes must be increased to retain the desired precision. For PIAAC, countries were asked to estimate the design effect of their design with such an increase in mind (guideline 4.3.2.B). However, the guideline was relaxed for this first cycle of PIAAC due to (1) uncertainties surrounding the quality of the design effect estimates produced using the Field Test data and (2) the limited amount of time available between the Field Test and the Main Study to allow changes to sample size goals of the survey.

Instead, countries with estimated large design effects were asked to modify their design to the extent possible to reduce the clustering of the sample. To compute the initial sample size, countries were allowed to use a design effect of 1.50 (if the expected design effect was greater than 1.50). However, countries are asked to report their best estimate of the design effect so that improvements to clustering and stratification may be identified for future cycles of PIAAC.

Section 14.5.2 contains information about the various expected eligibility rates used in the computation of the initial sample sizes by the participating countries and the plans for selecting reserve samples in case observed rates were different from the expected ones.

14.5.1 PIAAC target sample sizes

The minimum sample size requirements for the Main Study for the standard target population speaking the main language of the country was dependent on the optional components of the psychometric assessments administered in the country:

- Both problem solving and reading components – 5,000 minimum completes
- Problem solving only – 5,000 minimum completes
- Reading only – 4,500 minimum completes
- No optional components – 4,500 minimum completes

The definition of a completed case is given in TSG 4.3.3 as follows:

'Standard 4.3.3 A completed case is one that contains at least the following:

- *Responses to key background questions, including age, gender, highest level of schooling and employment status; and*
- *A completed Core instrument (i.e. the interviewer asked the respondent all Core questions or the Core instrument was not completed for a literacy-related reason [e.g. because of a language difficulty] or because the respondent was unable to read or write in any of a country's PIAAC official languages); or*

- *Responses to age and gender for literacy-related nonrespondents to the BQ/JRA.'*

To obtain a self-weighting standard design, the number of assessments in any other language had to be proportional to the number of people speaking the additional languages in the country. Countries that planned to report on general proficiency, regardless of the languages tested, had to achieve the appropriate minimum completed sample size shown above for their main language. Thus, the minimum sample size requirement for an individual country not only depended on the optional psychometric assessments administered and the number of languages being tested but also the number of reporting languages determined by the country.

Most countries conducted both the reading and problem-solving components. Cyprus,[18] Italy and Spain conducted the reading components only; Finland, Japan and the Russian Federation[19] conducted the problem-solving component only. France declined both optional assessments. Five countries performed the assessment in multiple languages. Canada, Estonia, Finland and the Slovak Republic conducted assessments in two languages; Spain conducted the assessment in five languages. The full list of the optional components of the psychometric assessment being conducted by the countries, including the languages of the assessments and the resulting required number of assessments, is presented in Table 14-19, and target sample sizes are given in Table 14-20 below.

[18] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[19] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 14-19: Required sample sizes by assessment language

Country	Assessment language and proportion of population speaking it (as available)	Optional components of psychometric assessment being conducted	Required sample size (general proficiency reporting in terms of language unless otherwise indicated)[1]
Australia	English	R, PS	5,000
Austria	German (88.5%)	R, PS	5,000
Canada	Canadian English (67.3%)	R, PS	5,000
	French (21.1)	R, PS	5,000
Cyprus [20]	Greek (84.1%)	R	4,500
Czech Republic	Czech	R, PS	5,000
Denmark	Danish (92%)	R, PS	5,000
England/N. Ireland (UK)	UK English	R, PS	5,000
	UK English	R, PS	5,000
Estonia	Estonian (67%)	R, PS	5,000
	Russian (33%)	R, PS	2,500
Finland	Finnish (90.5%)	PS	5,000
	Swedish (5%)	PS	276
Flanders (Belgium)	Dutch	R, PS	5,000
France	French	None	4,500
Germany	German	R, PS	5,000
Ireland	English	R, PS	5,000
Italy	Italian	R	4 500
Japan	Japanese (~100%)	PS	5,000
Korea	Korean	R, PS	5,000
Netherlands	Dutch	R, PS	5,000
Norway	Norwegian (Bokmål)	R, PS	5,000
Poland	Polish	R, PS	5,000
Russian Federation [21]	Russian (98.2%)	PS	5,000
Slovak Republic	Slovak (89.8%)	R, PS	5,000
	Hungarian (10.2%)	R, PS	568
Spain	Castellano (60%)	R	4,500
	Gallego (6%)	R	225
	Catalan (18%)	R	675
	Valencian (11%)	R	410
	Euskera (5%)	R	190
Sweden	Swedish	R, PS	5,000
United States	English (91.5%)	R, PS	5,000

[1] The required sample size in this table does not consider the occurrence of oversampling in some countries.

[20] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[21] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

14.5.2 Eligibility rates and reserve samples

The eligibility rate assumptions specified by countries were reviewed to help ensure that initial sample sizes were large enough to achieve the required number of assessments. Countries including a dwelling unit sample as part of their sample design were further required to provide an estimated screener eligibility rate. Selected units found to be vacant, for seasonal use only, not actually dwelling units, or without persons ages 16 to 65 were considered ineligible for the survey and had to be accounted for in the derivation of the final sample size.

The expected response rates reported during the National Survey Design and Planning Report process were taken into account to ensure that the initial samples sizes were large enough to yield the required number of assessments. Some adjustments to these expected rates were made based on Field Test experience.

It is difficult to predict the nonresponse and ineligibility rates for a survey like PIAAC. As a result, the Consortium encouraged each country to consider selecting a reserve sample of 10% or more of the size of the main initial (original) sample. The requirement was to select the reserve sample at the same time as the original sample and then set it aside and not use it unless sample monitoring showed potential for shortfall. Reserve samples were recommended over supplemental samples because computing the selection probabilities is simpler with a reserve sample than supplemental samples. The same concept was used if a country was concerned about exceeding the target sample size by a significant amount. After selecting a 110% sample, the country was able to release to the field a sample that was less than 100% by randomly selecting (subsetting) from the original sample and then releasing more sample as needed. Also the countries could split the reserve sample randomly into several “release” groups as long as the release group by itself was representative of the country (not any particular subgroup).

The target sample sizes for each stage, including the target person sample sizes, are presented in Table 14-20.

Table 14-20: Main Study target sample sizes

Country	Sample size				Target number of completes*	PIAAC standard**
	PSUs	SSUs	DUs	Persons		
Australia	2,136	2,136	14,423	11,250	9,000[1]	5,000
Austria				10,000	5,000	5,000
Canada [2]	217		49,234	34,464	25,267	10,000
Cyprus [22]			16,215	4,986	4,500	4,500
Czech Republic [3]	284	400	15,660	6,312	6,000	5,000
Denmark [4]				14 100	6 900	5,000
England (UK)	488	13,664	13,664	7,429	4,850	5,000
Estonia				13,000	7,500	7,500
Finland				8,000	5,300	5,276
Flanders (Belgium)				10,960	5,000	5,000
France	525			10,500	5,200	4,500
Germany	320			11,406	5,000	5,000
Ireland	700		13,600	8 092	6,200	5,000
Italy	260		17,520	7,742	4,500	4,500
Japan	459			13,000	5,000	5,000
Korea	883		8,330	7,296	5,000	5,000
Netherlands				10,256	5,000	5,000
Norway				9,453	5,000	5,000
Northern Ireland (UK)		9,470	9,470	5,143	3,492	5,000
Poland	85 urban 1,086 rural			13,430	9,132 [5]	5,000
Russian Federation[23]	25 [6]	93	9,630	5,540	5,000	5,000
Slovak Republic	562			9,280	5,568	5,568
Spain	1,200			14,400	6,000	6,000
Sweden				10,000	5,100	5,000
United States	80	901	9,610	6,371	5,000	5,000

(grey shading) indicates there is no such stage in the country's sample design.
*Targets include multiple languages and oversampling within target population, unless otherwise noted.
** Targets include multiple languages; there are no PIAAC standards for oversampled populations.
[1] 7,922 of the targeted completes were expected to be ages 16-65.
[2] Values include oversamples of 20,488 dwellings and 14,342 persons for 9,756 completes.
[3] Values include 5,923 sampled DUs, 1,052 sampled persons, and 1,000 targeted completes for the country-specific sample.
[4] Values do not include the Programme for International Student Assessment oversample, which was not part of the PIAAC sample.
[5] Includes oversample of 5,000 persons ages 19-26.
[6] Although the Russian Federation selected 25 PSUs, only 23 PSUs were included in the final analyses (Moscow and Moscow region were excluded due to data issues)

[22] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[23] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

14.6 Sample selection results

Table 14-21 provides the final sample sizes for each stage of sampling for each country. Table 16-7 provides the final number of respondents (with a final sampling weight).

Table 14-21: Main Study selected sample sizes by sampling stage

Country	Sample size			
	PSUs	SSUs	DUs	Persons
Australia	~2,200	~2,200	14,634	9,725 [1]
Austria				10,000
Canada	217		49,487	33,987
Cyprus [24]			8,514	5,095
Czech Republic	284	400	17,069	6,907
Denmark				16,040
England (UK)	488	13,664	13,664	7,933
Estonia				13,000
Finland				8,099
Flanders (Belgium)				9,200
France	525			10,500
Germany	277			10,240
Ireland	700		10,500	6,442
Italy	260		11,592	7,377
Japan	459			11,000
Korea	883		8,330	7,296
Netherlands				10,256
Northern Ireland (UK)		9,480	9,480	4,937
Norway				8,506
Poland	85 urban 1,086 rural			18 774
Russian Federation[25]	25 [2]	93	9,376	4,199
Slovak Republic	562			9,280
Spain	1,200			14,400
Sweden				10,000
United States	80	896	9,468	6,100

[grey box] indicates that there is no such stage in the country's sample design.
[1] 8,433 were ages 16-65.
[2] Although the Russian Federation selected 25 PSUs, only 23 PSUs were included in the final analyses (Moscow and Moscow region were excluded due to data issues)

14.7 Sampling quality control checks

The Consortium developed a comprehensive set of quality assurance and quality control checks to ensure PIAAC produced high-quality data that were comparable across countries. Section 16.1 contains a description of the quality assurance and quality control procedures developed for all sampling activities, including sample design and selection results. Countries were required to

[24] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[25] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

complete quality control sample selection forms, which collected sampling information for each stage of selection using standard templates. The templates were designed to capture aggregated information that was necessary for verifying that the sample was representative of the target population and that sampling was conducted in an unbiased and randomized way. For example, at each stage countries were asked to estimate and report the total target population within each stratum so that distributions by stratum could be reviewed at each sampling stage. The Consortium carried out all sampling quality control checks as listed in section 16.1 and informed the countries of the approval of their plans/procedures or asked for revisions to aspects that did not meet the PIAAC standards.

Table 14-22 provides a summary of the sample design and selection quality assessment. For the sampling plan, it was essential that a complete sampling plan was provided, and that the country responded to feedback from the Consortium. For the sampling plan, a cautionary remark was given to the Russian Federation[26] due to an insufficient number of PSUs selected. As it relates to the sample selection process conducted in the country's home office, it was important that complete QC sample selection forms were provided prior to data collection, that each person in the PIAAC target population had a nonzero and known (calculable) probability of selection resulting from the application of established and professionally recognized principles of scientific sampling, and that there was no substitution of sampling units. As indicated in Table 14-22, cautionary remarks were given to Australia (quality level unknown due to country confidentiality restrictions or unavailability of data), Czech Republic (for late sample selection forms), Germany (for simulated probabilities of selection), the Russian Federation[27] (noncompliance in completing the quality control forms), and Japan (for an approved deviation of the TSG, given the disastrous earthquake. The design accounted for the affected PSUs through combining strata, increasing sample sizes in affected strata, and using weighting procedures to reduce bias), With regard to sample selection processes that were conducted in the field, countries were assessed according to the following criteria ensuring that:

- persons were selected from within households using a fully enumerated grid of household members,
- each person in the PIAAC target population had a nonzero and known (calculable) probability of selection resulting from the application of established and professionally recognized principles of scientific sampling,
- no more than two persons were selected in a household,
- less than 10% of households had two persons selected, and
- there was no substitution of sampling units.

Only cautionary remarks were given to Australia (quality level unknown due to country confidentiality restrictions or unavailability of data) and the UK (imputed theoretical person base weights for 52 cases (49 in England and three in Northern Ireland) due to a technical problem with the contact data that the interviewers entered).

[26] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[27] Please refer to the above note regarding the Russian Federation.

Table 14-22: PIAAC sample design and selection outcome summary

	Sampling Plan	Sample Selection	
		Home Office	In Field
Australia	P	C-U	C-U
Austria	P	P	N/A
Flanders (Belgium)	P	P	N/A
Canada	P	P	P
Cyprus [28]	P	P	P
Czech Republic	P	C-NC	P
Denmark	P	P	N/A
England (UK)	P	P	C-PC
Estonia	P	P	N/A
Finland	P	P	N/A
Germany	P	C	N/A
Ireland	P	P	P
Italy	P	P	P
Japan	P	C-A	N/A
Korea	P	P	P
Netherlands	P	P	N/A
Northern Ireland (UK)	P	P	C-PC
Norway	P	P	N/A
Poland	P	P	N/A
Russian Federation[29]	C-PC	C-NC	P
Slovak Republic	P	P	N/A
Spain	P	P	N/A
Sweden	P	P	N/A
United States	P	P	P

P: Pass (relevant requirement completely met)
C: Caution (relevant requirement met to a reasonable extent)
C-A: Caution, approved deviation
C-NC: Caution, did not comply
C-PC: Caution, partial compliance
C-U: Caution, quality level unknown due to country confidentiality restrictions or unavailability of data

14.8 Respondent incentives

Respondent incentives have been shown to be effective for improving response rates without affecting the respondent's performance. As a result, the use of incentives can potentially reduce bias in the estimates. As such, countries were permitted to offer modest incentives to obtain respondent cooperation, such as a monetary or nonmonetary incentive (e.g., pen, notepad, candy, mug, voucher, gift certificate). A variety of incentives were offered across the participating countries with the exception of two countries: Australia and Canada have rules preventing the

[28] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[29] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

use of incentives in government surveys. Table 6-8 contains the type of incentives used during the Main Study data collection in PIAAC.

14.9 Recommendations for future cycles

Based on the Field Test and Main Study experience of PIAAC Round 1, the Consortium is proposing a series of recommendations for future cycles of PIAAC as it relates to sampling activities.

1. Countries should follow the TSG on the qualifications of the National Sampling Manager.
2. The Consortium and countries should work together to provide the BQ in as many languages as possible so that background information can be used in the generation of plausible values in case the person speaks a different language than the assessment language(s) offered.
3. Countries should evaluate the quality of the frames from the start so they have adequate time to look for alternatives if the quality (and coverage) of the frame does not meet the standards.
4. Before countries move forward with the sample that has been selected, the QC sample selection forms must be reviewed by the Consortium, with feedback provided.
5. Before countries submit sample monitoring forms, all numbers should be double checked. The Consortium will insert some automated checks into the forms to help ensure the forms are completed accurately.
6. Countries should use the Response Rate Toolkit to compute the response rates for the forms, or to check any automated program that was developed.
7. Countries should use the results of PIAAC to improve upon the stratification and sorting scheme. The nonresponse bias analysis and the scores can be used to identify better stratification and sorting variables, such as education, employment and other variables that are correlated with the scores.
8. Countries should use the design effects to identify ways to improve the sample design. That is, countries should evaluate how to reduce the clustering and unequal probabilities effects as plans occur for the next cycle.
9. While preparing plans for the next cycle, initial sample sizes should take into account the impact of the design components (cluster sizes, stratification, variation in weights, multiple imputation) on the resulting DEFFs observed in Cycle 1 (or an expected DEFF due to design improvements since Cycle 1) so that the quality of the resulting estimates is comparable across countries. Countries should plan to increase their sample sizes to account for the large design effects to arrive at an acceptable effective sample size, or make changes in their sample designs to reduce design effects.
10. Countries need to follow the schedules of all QC sampling activities so there is adequate time to identify problems and to incorporate changes to correct mistakes in a timely fashion.

References

Cox, L. (1987). A constructive procedure for unbiased controlled rounding. *Journal of the American Statistical Association*, *82*(398), 520-524.

Deville, J., & Tillé, Y. (2004). Efficient balanced sampling: The cube method. *Biometrika*, *91*(4), 893-912.

Tillé, Y. (2006). *Sampling algorithms*. New York, NY: Springer.

Chapter 15: Survey Weighting and Variance Estimation

Leyla Mohadjer, Tom Krenzke, Wendy Van de Kerckhove and Valerie Hsu, Westat

This chapter describes the methods that countries used to compute sampling weights and estimate variances through the use of replicate weights. The purpose of calculating sampling weights for PIAAC is to permit inferences from persons included in the sample to the population from which they were drawn and to have the tabulations reflect estimates of the population totals. Sampling weights can be considered as estimated measures of the number of units in the target population that a sampled case represents. Weighting incorporates several features of the survey, including the probabilities of selection of units in the sample and adjustments for nonresponse and any known differences between the selected sample and the total target population. Differences between the sample and the population may arise because of sampling variability, differential response rates or coverage rates among subgroups of the population, and other types of response errors, such as misclassification errors.

In PIAAC, survey weighting was performed to accomplish the following objectives:

- To permit unbiased estimates by compensating for possible disproportionate sampling of various subgroups in the sample
- To minimize biases arising from differences between respondents and nonrespondents
- To compensate for noncoverage in the sample due to inadequacies in the sampling frame or other reasons for noncoverage
- To bring data up to the dimensions of the population totals
- To reduce sampling errors by using auxiliary data on population characteristics that are known with a high degree of accuracy
- To facilitate the estimation of variances through the use of the replication approach

15.1 Survey weighting

Weighting involves designing adjustment factors to compensate for variable probabilities of selection and to reduce potential bias due to nonresponse, deficiencies in the sampling frame and other complications that may arise during the sample selection process. This section provides a description of the standard weighting steps employed in the first round of PIAAC. Countries were required to follow the weighting process outlined in the PIAAC Weighting and Variance Estimation Plan produced by the Consortium, which followed the standards and guidelines in Section 14 of the PIAAC Technical Standards and Guidelines. It described the weighting process, including the weighting steps, treatment of different disposition codes, calculation of weighting adjustment factors, assignment of variance strata and variance units, and creation of replicate weights. Using the weighting approach described in the Weighting and Variance

Estimation Plan for all countries ensured comparable estimates of proficiency and their sampling error across countries.

A final weight is required for all sampled persons with a completed BQ and BQ literacy-related nonrespondents (LRNRs) with age and gender collected. The BQ LRNRs with age and gender collected receive a final weight despite the lack of BQ or assessment data because they are considered part of the PIAAC target population and cannot be represented by survey respondents (see section 15.1.3). There were a number of steps in the development of the final weights intended for use in the estimation and analysis:

1. Assignment of a household base weight to each sampled household to compensate for differential probabilities of selection (for screener countries[1] only)
2. Household-level eligibility and nonresponse adjustments to reduce potential biases arising from differences between respondents and nonrespondents (for screener countries only)
3. Assignment of a person base weight to each sampled person to compensate for differential probabilities of selection
4. Person-level eligibility adjustment (for registry countries[2] only) and nonresponse adjustments
5. Trimming to reduce the impact of large weights, if necessary
6. Calibration of the person weights to independent control totals to compensate for noncoverage in the sample due to deficiencies in the sampling frame

The succeeding sections describe each of the weighting steps in detail. A summary of the adjustment factors and resulting weights at each weighting step is provided in Tables 15-1A and 15-1B for registry and screener countries, respectively.

[1] Screener countries refer to countries whose sample design included a screener stage.

[2] Registry countries refer to countries whose sample design did not include a screener stage.

Table 15-1A: Adjustment factors and weights for registry countries

Weighting Step	Factor	Weight
Base weight	N/A	$W_l = \dfrac{1}{P_{hl}}$
Unknown eligibility adjustment	$F_{1l} = \begin{cases} \dfrac{S_R + S_{NR} + S_{L1} + S_{L2} + S_D + S_I + S_U}{S_R + S_{NR} + S_{L1} + S_{L2} + S_D + S_I} & \text{if } l \in \mathrm{I} \\ \dfrac{S_R + S_{NR} + S_{L1} + S_{L2} + S_D}{S_R + S_{NR} + S_{L1} + S_{L2} + S_D + S_I} & \text{if } l \in \mathrm{U} \\ 1 & \text{if } l \in \mathrm{R, NR, L1, L2, D} \end{cases}$	$W_l F_{1l}$
Nonliteracy-related nonresponse adjustment	$F_{3l} = \begin{cases} 1 & \text{if } l \in \mathrm{L1, L2, I} \\ \dfrac{S_R + S_{NR} + S_D + S_U}{S_R} & \text{if } l \in \mathrm{R} \\ 0 & \text{if } l \in \mathrm{NR, D, U} \end{cases}$	$W_l F_{1l} F_{3l}$
Literacy-related nonresponse adjustment	$F_{4l} = \begin{cases} 1 & \text{if } l \in \mathrm{R, I} \\ \dfrac{S_{L1} + S_{L2}}{S_{L1}} & \text{if } l \in \mathrm{L1} \\ 0 & \text{if } l \in \mathrm{L2} \end{cases}$	$W_l F_{1l} F_{3l} F_{4l}$
Trimming*	$F_{5l} = \begin{cases} 1 & \text{if } W_l F_{1l} F_{3l} F_{4l} \le cutoff \\ \dfrac{cutoff}{W_l F_{1l} F_{3l} F_{4l}} & \text{if } W_l F_{1l} F_{3l} F_{4l} > cutoff \end{cases}$	$W_l F_{1l} F_{3l} F_{4l} F_{5l}$
Calibration	$F_{6l} = \dfrac{S^*}{S_R + S_{L1}}$ (for post-stratification) See Deming and Stephan (1940) for raking adjustments and Särndal, Swenson, and Wretman (1992) for GREG estimation.	$W_l F_{1l} F_{3l} F_{4l} F_{5l} F_{6l}$

* If the Consortium computed the sampling weights, an initial calibration step was performed prior to trimming (i.e., one iteration of calibration, trimming (if necessary), and recalibration was performed following the nonresponse adjustments).

Note: The factors and weights shown here are for a person *l*. The persons can be classified as R: BQ respondent who is not assessment literacy-related nonrespondent, L1: BQ literacy-related nonrespondent with age and gender successfully collected or assessment literacy-related nonrespondent, L2: BQ literacy-related nonrespondent with age or gender not successfully collected, NR: BQ nonliteracy-related nonrespondent, I: ineligible, D: sampled person with a disability, or U: sampled person with unknown eligibility status. *S* represents the sum of the prior-stage weights over records in the same adjustment cell as person *l*, and *S** is the control total for the cell. *P* represents the selection probability. The factor F_2 is reserved for countries with screeners.

Table 15-1B: Adjustment factors and weights for screener countries

Stage	Weighting Step	Factor	Weight
Screener	Base weight	N/A	$W_k = \frac{1}{P_{hi} CP_{hik}}$
	Unknown eligibility a $W_l F_{3l} F_{4l} F_{5l} F_{6l}$ djustment	$F_{1k} = \begin{cases} \frac{S_L + S_R + S_{NR} + S_I + S_U}{S_L + S_R + S_{NR} + S_I} & \text{if } k \in \mathrm{I} \\ \frac{S_L + S_R + S_{NR}}{S_L + S_R + S_{NR} + S_I} & \text{if } k \in \mathrm{U} \\ 1 & \text{if } k \in \mathrm{L, R, NR} \end{cases}$	$W_k F_{1k}$
	Nonresponse adjustment	$F_{2k} = \begin{cases} 1 & \text{if } k \in \mathrm{L, I} \\ \frac{S_R + S_{NR} + S_U}{S_R} & \text{if } k \in \mathrm{R} \\ 0 & \text{if } k \in \mathrm{NR, U} \end{cases}$	$W_k F_{1k} F_{2k}$
BQ	Base weight	N/A	$W_l = W_k F_{1k} F_{2k} \frac{1}{CP_{h}}$
	Nonliteracy-related Nonresponse adjustment	$F_{3l} = \begin{cases} 1 & \text{if } l \in \mathrm{L, I} \\ \frac{S_R + S_{NR} + S_D}{S_R} & \text{if } l \in \mathrm{R} \\ 0 & \text{if } l \in \mathrm{NR, D} \end{cases}$	$W_l F_{3l}$
	Literacy-related nonresponse adjustment	$F_{4l} = \begin{cases} 1 & \text{if } l \notin \mathrm{L} \\ \frac{S'^{BQ}_L + S'^{MAIN}_L + S'^{SCR}_L}{S'^{BQ}_L + S'^{MAIN}_L} & \text{if } l \in \mathrm{L^{BQ}} \text{ or } \mathrm{L^{MAIN}} \\ 0 & \text{if } l \in \mathrm{L^{SCR}} \end{cases}$	$W_l F_{3l} F_{4l}$
	Trimming*	$F_{5l} = \begin{cases} 1 & \text{if } W_l F_{3l} F_{4l} \le cutoff \\ \frac{cutoff}{W_l F_{3l} F_{4l}} & \text{if } W_l F_{3l} F_{4l} > cutoff \end{cases}$	$W_l F_{3l} F_{4l} F_{5l}$
	Calibration	$F_{6l} = \frac{S*}{S_R + S_{L^{BQ}} + S_{L^{MAIN}}}$ (for post-stratification) See Deming and Stephan (1940) for raking adjustments and Särndal, Swenson, and Wretman (1992) for GREG estimation.	

* If the Consortium computed the sampling weights, an initial calibration step was performed prior to trimming (i.e., one iteration of calibration, trimming (if necessary), and recalibration was performed following the nonresponse adjustments).

NOTE: The factors and weights shown here are for a household *k* or person *l*. The households and persons can be classified as R: respondent, L: literacy-related nonrespondent, NR: nonliteracy-related nonrespondent, I: ineligible, D: sampled person with a disability, or U: unknown eligibility. *S* represents the sum of the prior-stage weights over records in the same adjustment cell as household *k* or person *l*, *S'* is the sum of screener base weights, and *S** is the control total for the cell. *P* represents the selection probability.

15.1.1 Preliminary steps in weighting

Countries were responsible for selecting the variables that were used in their nonresponse and calibration weighting adjustments. Prior to weighting, countries were required to evaluate the variables being considered for the weighting adjustments in their PIAAC main sample.

For the nonresponse adjustment, variables needed to be available for all eligible units and be related to proficiency and response propensity. The pool of potential nonresponse adjustment variables came from the sampling frame (and/or the screener) or other external sources. A common source of nonresponse adjustment variables for screener countries was a country census. For registry countries, the registry data were highly beneficial during the nonresponse adjustment.

For the calibration adjustment, all variables selected by countries were required to have reliable control totals and be available for all BQ respondents and LRNRs with age and gender collected. The quality of the data from the external sources had to exceed the quality of data from PIAAC (e.g., the mean square errors of the external estimates needed to be smaller than those of the uncalibrated estimates from the survey). The concepts, definitions and coverage of the data (counts) from the external sources needed to be the same as those employed by PIAAC. Additionally, the year of the control totals needed to be as close to the data collection period as possible, ideally covering the same time period as the field period.

Variables used for nonresponse adjustment and in calibration must have less than 5% missing data. If the amount of missing data of the variables used in weighting adjustments did not exceed the 5% threshold, countries were required to follow the weighting standards and guidelines on imputing for missing data.

15.1.2 Household-level weighting adjustments

This section outlines the weighting process at the household level for screener countries, which included the creation of the household base weights that reflected the household selection probability and was adjusted for unknown eligibility and nonresponse to the screener.

Household base weights

For screener countries, the household base weight was assigned to all sampled households and was computed as the reciprocal of the household selection probability. For screener countries with a multistage sample design, the household selection probability corresponded to the product of the conditional selection probabilities at each stage. For example, if households were selected within primary sampling units (PSUs), then the household base weight would be

$$W_k = \frac{1}{P_{hi} CP_{hik}},$$

where P_{hi} is the probability of selecting PSU i in stratum h, and CP_{hik} is the conditional probability of selecting household k within PSU i of stratum h.

The household selection probability also reflected any duplicate records in the sampling frame or any changes to the subsampling procedures.

Household unknown eligibility adjustment

Before any household-level nonresponse adjustment was applied, an adjustment for unknown eligibility was performed if the eligibility status of some households could not be determined. In this step, a portion of the weights of the households with unknown eligibility status (i.e., whether they contained a person age 16 to 65) was distributed to ineligible cases. An adjustment factor was computed as the proportion eligible among those with known eligibility status to down-weight the cases with unknown eligibility status (accounting for an estimated proportion that was ineligible). The down-weighted unknown eligibility cases were then treated as eligible nonrespondents. This adjustment was done within weighting cells defined for the unknown eligibility adjustment (see Table 15-3).

Household nonresponse adjustment

For the screener nonresponse adjustment, the nonrespondents were divided into two categories. The first consisted of cases involving nonliteracy-related nonresponse. Examples of this category included refusals and nonresponse due to speech impairment. Nonliteracy-related nonrespondents were likely to be similar to respondents with respect to proficiency scores. The second category was literacy-related nonresponse. Language problem was the only type of literacy-related nonresponse at the screener level. Households with this type of nonresponse were presumed to differ from responding households with respect to proficiency. Therefore, the weighting procedures adjusted the weights of the respondents to represent the nonliteracy-related nonrespondents only. The weights of the LRNRs were not adjusted during the screener-level nonresponse adjustment because their proficiency was expected to differ from that of respondents. The contribution of the screener level literacy-related nonresponse to the total population was accounted for by the literacy-related nonresponse adjustment carried out at the person level involving the assessment LRNRs (see section 15.1.3).

The next step in the weighting process was to adjust the unknown eligibility-adjusted weights to reduce potential bias as a result of nonresponse to the screener. An adjustment was made to distribute the screener unknown eligibility-adjusted weights of the nonliteracy-related nonrespondents to the screener respondents. The nonresponse adjustment was performed within cells that were defined based on pre-selected weighting variables that were found to be related to proficiency and to response propensity (see Table 15-3). Within each adjustment cell, the household unknown eligibility-adjusted weights of nonrespondents were redistributed over a relatively large pool of cases (approximately 30 or more respondents). Additionally, the amount of variation in the nonresponse adjustment factors was kept to a minimum by limiting the maximum allowable nonresponse adjustment factor, which was a function of the achieved screener response rate.

15.1.3 Person-level weighting adjustments

This section describes the process of creating the person-level weights, including the computation of person base weights; the person unknown eligibility adjustment that applied to registry countries only; the nonresponse adjustment procedure designed to reduce potential nonresponse bias; the calibration of weights to control totals; and the general trimming procedure used to reduce the impact of extreme weights.

Person base weights

For screener countries, the person base weights accounted for both nonresponse to the household screener and differential within-household selection rates. The person base weights were computed as the product of the household nonresponse-adjusted weight and the reciprocal of the within-household person selection probability.

For registry countries, the base weight for each sampled person was computed as the reciprocal of the person selection probability.

Person unknown eligibility adjustment

For registry countries, an adjustment for person unknown eligibility was performed if the eligibility status of some sampled persons could not be determined due to the inability of the survey to locate and interview these selected persons not residing at the address listed in the registry (see section 16.2.2 for a discussion on inaccessible sampled persons). In the person unknown eligibility adjustment, a portion of the person base weights of the sampled persons with unknown eligibility status was distributed to the ineligible cases. An adjustment factor was computed as the proportion eligible among those with known eligibility status to down-weight the cases with unknown eligibility status (accounting for an estimated proportion that was ineligible). The down-weighted unknown eligibility cases were then treated as eligible nonrespondents in the nonresponse adjustment.

Person nonliteracy-related nonresponse adjustment

For the nonresponse adjustment, the nonrespondents were divided into two categories. The first category consisted of nonliteracy-related nonrespondents (e.g., refusals and inaccessibles with known eligibility) and sampled persons with a disability (e.g., hearing impairment and physical disability). They were likely to be similar to respondents with respect to proficiency scores. The second category was literacy-related nonresponse (LRNR). Types of literacy-related nonresponse include language problem, reading and writing difficulty, and learning-mental disability. Sampled persons with this type of nonresponse were presumed to differ from respondents with respect to proficiency. Therefore, LRNRs received a different treatment than nonliteracy-related nonrespondents.

As mentioned earlier, for screener countries, an adjustment was made to distribute the person base weights of the nonliteracy-related nonrespondents and sampled persons with a disability to the respondents' weights.

For registry countries, excluded inaccessible sampled persons were treated as nonliteracy-related nonrespondents in weighting. An adjustment was made to distribute the person unknown eligibility-adjusted weights of the nonliteracy-related nonrespondents, sampled persons with a disability, and down-weighted unknown eligibility cases to respondents.

The nonresponse adjustment was performed within cells that were defined based on pre-selected weighting variables that were found to be related to proficiency and to response propensity (see Table 15-3). Within each adjustment cell, the person unknown eligibility-adjusted weights of nonrespondents were redistributed over a relatively large pool of cases (approximately 30 or more respondents). Additionally, the amount of variation in the nonresponse adjustment factors was kept to a minimum by limiting the maximum allowable nonresponse adjustment factor, which depended on the achieved BQ response rate.

Person literacy-related nonresponse adjustment

For screener countries, the weights of the BQ and assessment LRNRs were adjusted to account for the screener LRNRs. This adjustment was necessary primarily to allow both the BQ and assessment LRNRs to represent the screener LRNRs in the calibration procedure. This adjustment assumed that the LRNRs to the screener, BQ and assessment were similar in proficiency.

For registry countries, the weights of the BQ LRNRs with age and gender collected and assessment LRNRs were adjusted to account for the weights of the BQ LRNRs without age and gender collected.

Involving the assessment LRNRs in the literacy-related nonresponse adjustment offered several advantages. This approach (1) reduced the mean square error in the resulting estimates, (2) provided stability in the weight adjustment and reduced the variations in the weights and in the estimates, (3) reduced bias under the assumption that the assessment LRNRs were more similar to the BQ LRNRs than the BQ nonliteracy-related nonrespondents, and 4) addressed the issue that sampled persons may or may not have completed the BQ because of an arbitrary reason (e.g., unavailable bilingual interviewer or interpreter).

Calibration

To address undercoverage bias, to reduce the mean square error of estimates and to create consistency with statistics from other studies, the next weighting step was to adjust the survey weights to match population control totals. At minimum, weights were benchmarked to control totals for age and gender. Respondents who completed the BQ and BQ LRNRs received a final weight and were included in calibration. If the Consortium performed the weighting adjustments, one iteration of calibration, trimming (if necessary) and recalibration was performed following the nonresponse adjustments. Not all countries that performed their own weighting included the initial calibration prior to trimming.

Three main calibration techniques employed by countries are post-stratification, raking and generalized regression estimators (GREG). Post-stratification adjusts survey weights of respondents so that the weighted sample distribution is the same as some known population distribution (i.e., the sums of the adjusted weights of the respondents are equal to known population totals for certain subgroups of the population). The raking procedure uses an iterative procedure to adjust the survey estimates to the known marginal totals of several categorical variables. The GREG estimator is a model-assisted approach that can be used to adjust weights to exploit explicitly the relationship between a survey variable and auxiliary variables.

Trimming the outliers

Even a carefully designed sample could not fully prevent the need for reducing extreme weights. Sample designs that included the selection of dwelling units had more variability in the weights compared to directly sampling persons from registries because of unequal household sizes. The use of nonresponse and calibration adjustments also introduced variations in sampling weights.

Weight trimming introduced some bias into the sampling weights. However, the trimming adjustment in most cases reduced the sampling error component of the overall mean square error more than it increased the bias as the adjustment was applied to only a relatively small number of weights (Lee, 1995).

The person weights were trimmed as necessary after the first calibration. Using a design-based procedure, cells for trimming were formed from groups that were expected to be approximately self-weighting. In each cell, weights above a cutoff value were trimmed down to the designated cutoff. To define the trimming cut point, the Consortium examined the coefficient of variation (CV) based on the weights after raking (the cut point was calculated separately by domain in case oversampling was used for some domains). The Consortium trimmed the weights that were over $3.5 \times \sqrt{1 + CV^2}$ times the median raked weight (within each trimming cell, if sampling rates varied by sampling domains). In a few instances, a review of the distribution of the raked weights revealed that a different cut point was more appropriate. Some countries that performed their own weighting used different criteria for trimming. During trimming, the trimming factor was applied to each replicate weight. After trimming, the weights were recalibrated back to the control totals.

15.1.4 Weighting quality control checks

Quality control (QC) checks were performed for both the full sample and replicate weights after each adjustment in the weighting procedure to ensure proper implementation. The Consortium developed a battery of QC checks to review the weighting process for adherence to the weighting standards and guidelines and to check weight calculations for reasonableness and accuracy. Performing the weighting QC checks was essential for verifying that the final weights produced for estimation are appropriate (see section 16.1). The PIAAC schedule required the weighting QC checks to be conducted prior to the development of proficiency scores. Further checks were conducted after derivation of the proficiency scores if analyses showed any need for re-verification/correction of the weights.

15.1.5 Summary of country-specific weighting implementation

This section presents the weighting steps performed by countries, variables selected by countries for weighting adjustments and country-specific deviations from the weighting standards. All participating countries in PIAAC were responsible for selecting weighting variables and preparing files for weighting. The Consortium was responsible for deriving sampling weights for the Main Study for all countries. Countries that opted to compute their own weights were required to follow the standards and guidelines in Chapter 14 of the PIAAC Technical Standards and Guidelines and the PIAAC Weighting and Variance Estimation Plan. The weighting procedures described in the standards ensured that the estimates represent each country’s target population and reduce the potential for bias due to nonresponse.

Weighting steps performed by countries

Table 15-2 indicates each participating country’s weighting responsibility, sample design, weighting steps performed, and calibration method. Any deviations from the weighting standards and special weighting adjustments are noted in Table 15-5.

Table 15-2: Weighting steps, by country

Country	Weighting Respon-sibility	Design	Screener: Base Weight	Screener: Unknown Eligibility Adjustment	Screener: Nonresponse Adjustment	Background Questionnaire: Base Weight	Background Questionnaire: Unknown Eligibility Adjustment[1] *	Background Questionnaire: Nonresponse Adjustment (nonliteracy-related)	Background Questionnaire: Nonresponse Adjustment (literacy-related)[2]	Background Questionnaire: Trimming[3]	Background Questionnaire: Calibration
Australia	Country	Screener	Y	N	N	Y		Y	Y	N	GREG
Austria	Westat	Registry				Y	Y	Y	Y	Y	Raking
Canada	Country	Screener	Y	Y	Y	Y		Y	Y	Y	Raking
Cyprus[3]	Westat	Screener	Y	Y	Y	Y		Y	Y	Y	Raking
Czech Republic	Westat	Screener	Y	Y	Y	Y		Y	Y	Y	Raking
Denmark	Country	Registry				Y	NA	Y	Y	N	GREG
England (UK)	Westat	Screener	Y	Y	Y	Y		Y	Y[4]	Y	Raking
Estonia	Westat	Registry				Y	Y	Y	NA	Y	Raking
Finland	Country	Registry				Y	Y	Y	N	Y	GREG
Flanders (Belgium)	Westat	Registry				Y	NA	Y	NA	Y	Raking
France	Westat	Registry				Y	Y	Y	NA	N	Raking
Germany	Westat	Registry				Y	Y	Y	Y	Y	PS
Ireland	Westat	Screener	Y	Y	Y	Y		Y	Y	Y	Raking
Italy	Country	Screener	Y	Y	Y	Y	Y	Y	Y	Y	Raking
Japan	Country	Registry				Y	Y	Y	NA	Y	GREG
Korea	Westat	Screener	Y	Y	Y	Y		Y	Y	Y	Raking
Netherlands	Country	Registry				Y	Y	Y	Y	N	GREG
N. Ireland (UK)	Westat	Screener	Y	Y	Y	Y		Y	Y[4]	Y	Raking
Norway	Country	Registry				Y	Y	Y	Y	Y	Raking
Poland	Westat	Registry				Y	Y	Y	N	Y	Raking
Russian Federation[4]	Westat	Screener	Y	Y	Y	Y		NA	NA	Y	Raking
Slovak Republic	Westat	Registry				Y	NA	Y	Y	Y	Raking
Spain	Country	Registry				Y	Y	Y	NA	Y	GREG
Sweden	Country	Registry				Y	Y	N	Y	N	GREG
United States	Country	Screener	Y	Y	Y	Y		Y	Y	Y	Raking

▯: not applicable, Y: weighting step performed, N: weighting step not performed, NA: weighting step not needed, PS: post-stratification

[1*] NA: There were no cases with unknown eligibility status (i.e., DISP_CIBQ=24 and EXCFLG=2).

[2] NA: There were no LRNRs with age and gender not collected (i.e., DISP_CIBQ = 7, 8, or 9 and QCFLAG_LR = 2) or no LRNRs at the screener level (DISP_SCR=7).

[3] A value of "Y" indicates that the weighting process included a step to evaluate whether there were any extreme weights and trim if necessary. It does not indicate the outcome of the trimming (i.e., whether any weights were trimmed).

[4] In addition to the standard literacy-related nonresponse adjustment, LRNRs with age and gender successfully collected represented those with age or gender not successfully collected.

---

[3] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[4] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Weighting variables selected by countries

After data collection and data editing, countries were to conduct an analysis to select variables for weighting adjustments that would be most effective in reducing nonresponse bias. At minimum, this analysis was to involve a classification tree or logistic regression to evaluate the relationship of response status to potential weighting variables.

The list of weighting variables selected by each country is given in Table 15-3. Of the countries that provided information, all used age and gender in calibration, as required in the PIAAC Technical Standards and Guidelines, and region was also used in all countries in either calibration or nonresponse adjustment. In addition, the majority of countries included in their weighting adjustments at least one variable related to education, employment status or nationality, which have been shown to be correlated with proficiency.

Benchmark control totals used by countries

Control totals used in the benchmarking process were required to have the same definition and coverage of the target population as PIAAC (noninstitutionalized adults who are between age 16 and 65, including citizens and noncitizens). If not, the counts from the external sources needed to be adjusted to make these comparable to the survey estimates. All variables selected for benchmarking must have reliable control totals available. The quality of data from external sources must have exceeded the quality of data from PIAAC (e.g., the standard errors, or more generally, the mean square error of the external estimates needed to be smaller than those of the nonbenchmarked estimates from the survey). Table 15-4 presents the control total variables used in calibration for each country, including its source and exclusions from the target population.

Table 15-3: Weighting variables, by country

Country	Screener Nonresponse Adjustment	Unknown Eligibility Adjustment	BQ Nonresponse Adjustment (nonliteracy-related)	BQ Nonresponse Adjustment (literacy-related)	Calibration
Australia	NA	NA	1 Cell	1 Cell	Highest educational attainment by state, labor force status by state by sex, labor force status by age group, state by part of state by sex by age group
Austria		Age by citizenship by education by urbanization (8)	Age by citizenship by education by urbanization (8)	Age by citizenship by education by urbanization (8)	Region by age (90), region by citizenship (18), region by level of urbanization by sex (48), sex by age by education (40)

Table 15-3 (cont.): Weighting variables, by country

Country	Screener Nonresponse Adjustment	Unknown Eligibility Adjustment	BQ Nonresponse Adjustment (nonliteracy-related)	BQ Nonresponse Adjustment (literacy-related)	Calibration
Canada	2011 Canadian Census short form (2A) questions and census paradata, 2006 census long form (2B) data at geographically aggregated level	?	The variables used for the screener NR adjustment were used. In addition, age and gender of the selected persons was used (333)	Delineation between general population and special subpopulations sample by province (30)	Age group and gender by province (130), educational attainment by province (52), immigration status and gender by province (21), aboriginal status and gender by province (24), census metropolitan area by province (26), linguistic minority status and gender by province (17)
Cyprus[6]	District by locale (7)	District by locale (9)	District, locale, age, education, gender (34)	District by locale (9)	Age by district (25), age by gender (10), age by education (15), gender by district (10), gender by education (6), language (2)

[6] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 15-3 (cont.): Weighting variables, by country

Country	Screener Nonresponse Adjustment	Unknown Eligibility Adjustment	BQ Nonresponse Adjustment (nonliteracy-related)	BQ Nonresponse Adjustment (literacy-related)	Calibration
Czech Republic	Region (8), municipality type (3), gender ratio quartiles (4), age ratio quartiles (4), employment status percentage quartiles (4), entrepreneurs percentage quartiles (4), education quartiles (4)	Region (8), municipality type (3), gender ratio quartiles (4), age ratio quartiles (4), employment status percentage quartiles (4), entrepreneurs percentage quartiles (4), education quartiles (4)	Municipality type (3), region (8), gender (2), age group (5), employment status percentage quartiles (4), entrepreneurs percentage quartiles (4), education quartiles (4)	1 Cell	Age by education (15), age by gender (10), education by gender (8), field of study by gender (16), work status by gender (14), region by employment status (24), region by education (32)
Denmark		NA	Income, region, education, type of family, mobility, marital status, socio-economic status, employment, gender (70)	1 Cell	Region (5), age (5), gender (20), immigration (4)

Table 15-3 (cont.): Weighting variables, by country

Country	Screener Nonresponse Adjustment	Unknown Eligibility Adjustment	BQ Nonresponse Adjustment (nonliteracy-related)	BQ Nonresponse Adjustment (literacy-related)	Calibration
England (UK)	Region (9), National Statistics 2001 Area Classification (21), index of multiple deprivation split into approximate deciles (10), 2001 census percentage living in social housing (9), 2001 census percentage Black or South Asian (7), 2001 census percentage of households that contain one person (10)	Region (9), National Statistics 2001 Area Classification (21), Index of multiple deprivation split into approximate deciles (10), 2001 census percentage living in social housing (9), 2001 census percentage Black or South Asian (7), 2001 census percentage of households that contain one person (10)	Region (9), national statistics 2001 area classification (21), index of multiple deprivation split into approximate deciles (10), 2001 census percentage living in social housing (9), 2001 census percentage Black or South Asian (7), 2001 census % of households that contain one person (10)	Region (9), national statistics 2001 area classification (21), index of multiple deprivation split into approximate deciles (10), 2001 census percentage living in social housing (9), 2001 census percentage Black or South Asian (7), 2001 census percentage of households that contain one person (10)	Gender by age (20), region (9), age by qualification s (17), gender by age by economic status (35)
Estonia		Age (2), gender (5), mother tongue (2), urbanization (3), county (15), percent of high education (4), percent of unemployment (4)	Age (2), gender (5), mother tongue (2), urbanization (3), county (15), percent of high education (4), percent of unemployment (4)	Age (2), gender (5), mother tongue (2), urbanization (3), county (15), percent of high education (4), percent of unemployment (4)	Gender by age (10), county (15), urbanization (3)
Finland		?	Gender (2), age (5), education (4), native language (3), region (5), urban/rural (3), family status (5)	Gender (2), age (5), education (4), native language (3), region (5), urban/rural (3), family status (5)	Gender (2), age (5), education (4), native language (3), region (5), urban/rural (3), family status (5)

Table 15-3 (cont.): Weighting variables, by country

Country	Screener Nonresponse Adjustment	Unknown Eligibility Adjustment	BQ Nonresponse Adjustment (nonliteracy-related)	BQ Nonresponse Adjustment (literacy-related)	Calibration
Flanders (Belgium)			Age (5), gender (2), province (5)		Age by work status (10), gender by work status (4)
France		Gender (2), age (5), region (3), income (5)	Gender (2), age (5), region (3), income (5)	Gender (2), age (5), region (3), income (5)	Age by gender (10), region (3), education (3), country of birth (2), employment status (3)
Germany[1]		Age, nationality, degree of urbanization	Age, nationality, degree of urbanization	1 Cell	Age, gender, region, education
Ireland	Percentage non-English language spoken at home (2), percentage unemployment (2), percentage with lower secondary-level education or below (2), owner occupied (2), regions (3)	Percentage non-English language spoken at home (2), percentage unemployment (2), percentage with lower secondary-level education or below (2), owner occupied (2), regions (3)	Gender (2), age (5), education (screener) (13)	Gender (2), age (5), education (screener) (13)	Region by age (40), region by gender (16), age by education (20), gender by education (8)

Table 15-3 (cont.): Weighting variables, by country

Country	Screener Nonresponse Adjustment	Unknown Eligibility Adjustment	BQ Nonresponse Adjustment (nonliteracy-related)	BQ Nonresponse Adjustment (literacy-related)	Calibration
Italy	Deciles of logit from model involving: Number of eligible persons in family, gender, age, municipality MOS, self-representing PSU indicator, region (10)	Quintiles of logit from model involving: Number of eligible persons in family, gender, age, municipality MOS, self-representing PSU indicator, region (5)	Number of eligible persons in family, gender, age, municipality MOS, self-representing PSU indicator, region (30)	1 Cell	Region by age (25), region by gender (10), region by education (15), region by employment (10)
Japan		Age (5), gender (2)	Age, gender, city size, region, type of building, area-level percentage: graduate from college, population density, household floor space, percentage of people employed in tertiary industry, number of persons per household, proportion of temporary workers to regular employees (20)	Age (5), gender (2)	Age (5), gender (2), education (6), employment status (3), geographic area (10)
Korea	Korea	Region (16), household type (3)	Region (16), household type (3)	Region (16), age (6), gender (2)	1 Cell

Table 15-3 (cont.): Weighting variables, by country

Country	Screener Nonresponse Adjustment	Unknown Eligibility Adjustment	BQ Nonresponse Adjustment (nonliteracy-related)	BQ Nonresponse Adjustment (literacy-related)	Calibration
Netherlands		Origin (3), household composition (5), social status (3), social status (3)	Origin (3), household composition (5), social status (3), social status (3)	Origin (3), household composition (5), social status (3)	Gender by age (10), origin by generation (5), group of provinces by degree of urbanization (18), household type (5), social status by income (25), term of registration in population registry (2), percentage of high level education by percentage of low level education (18)
Norway		?	Education, occupation, age group, industry and "special field"	?	Gender by age (10)

Table 15-3 (cont.): Weighting variables, by country

Country	Screener Nonresponse Adjustment	Unknown Eligibility Adjustment	BQ Nonresponse Adjustment (nonliteracy-related)	BQ Nonresponse Adjustment (literacy-related)	Calibration
Northern Ireland (UK)	Region (5), National Statistics 2001 Area Classification (20), 2001 census percentage living in social housing (9), index of multiple deprivation split into approximate deciles (10)	Region (5), National Statistics 2001 Area Classification (20), 2001 census percentage living in social housing (9), index of multiple deprivation split into approximate deciles (10)	Region (5), National Statistics 2001 Area Classification (20), 2001 census percentage living in social housing (9), index of multiple deprivation split into approximate deciles (10)	Region (5), National Statistics 2001 Area Classification (20), 2001 census percentage living in social housing (9), index of multiple deprivation split into approximate deciles (10)	Gender by age (20), region (5), age by qualifications (17), gender by age by economic status (35)
Poland		Income (4), age (5), population (9), region (16), number of cities per county (11), level of unemployment (5), proportion of middle-school students (4), computerization (4)	Income (4), age (5), population (9), region (16), number of cities per county (11), level of unemployment (5), proportion of middle-school students (4), computerization (4)	Income (4), age (5), population (9), region (16), number of cities per county (11), level of unemployment (5), proportion of middle-school students (4), computerization (4)	Gender by age (10), gender by region (32)

Table 15-3 (cont.): Weighting variables, by country

Country	Screener Nonresponse Adjustment	Unknown Eligibility Adjustment	BQ Nonresponse Adjustment (nonliteracy-related)	BQ Nonresponse Adjustment (literacy-related)	Calibration
Russian Federation[7]	Macro-region (8), type of settlement (3), type of district (3), education rate (3), unemployment rate (3)	Macro-region (8), type of settlement (3), type of district (3), education rate (3), unemployment rate (3)	NA	NA	Gender by age (20), education rate (3), macro-region (8)
Slovak Republic			Size of municipality (9), urban/rural (2), region (8), age by gender (10)	1 Cell	Size of municipality (9), urban/rural (2), region (8), age by gender (10)
Spain		Age (5), gender (2), nationality (2)	Age (5), gender (2), nationality (2), urbanicity (3), education (3), unemployment rate (4)	BQ LRNRs (1)	Gender (2), age (5), region (18), nationality (2), education (3)
Sweden		NA	NA	?	Education by sex by age (30), education by region (24), education by employment (9), education by income (12), education by country of birth (6)

[7] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 15-3 (cont.): Weighting variables, by country

Country	Screener Nonresponse Adjustment	Unknown Eligibility Adjustment	BQ Nonresponse Adjustment (nonliteracy-related)	BQ Nonresponse Adjustment (literacy-related)	Calibration
United States	Metropolitan Statistical Area, region, categorized percent of: Housing units occupied by owner, Hispanic or non-Hispanic Black, Hispanic, population age 18-64 unemployed, population below 150% of poverty, foreign born, household linguistically isolated, population age 25+ with high school education, population age 25+ with some college education, categorized household size (26)	Metropolitan Statistical Area, region, categorized percent of: Housing units occupied by owner, Hispanic or non-Hispanic Black, Hispanic, population age 18-64 unemployed, population below 150% of poverty, foreign born, household linguistically isolated, population age 25+ with high school education, population age 25+ with some college education, categorized household size (26)	Metropolitan Statistical Area, region, categorized percent of: Housing units occupied by owner, population age 25+ with at least high school education, Hispanic or non-Hispanic Black, Hispanic, population age 18-64 unemployed, foreign born, household linguistically isolated, population age 18-64 employed, population age 25+ with some college education, categorized household size, best age category (after imputation), indicator for children under age 16 in household, best gender, best race/ethnicity (after imputation) (26)	1 Cell	Educational attainment by race/ethnicity (12), education attainment by age (20), education attainment by gender (8), race/ethnicity by age (9), race/ethnicity by gender (6), country of birth by age (10), country of birth by region (8)

 not applicable, NA: weighting step not performed, ?: unknown/received no information from country

[1] The number of categories is not provided for confidentiality reasons.

NOTE: Numbers in parentheses indicate the number of categories.

Table 15-4: Benchmark control totals, by country

Country	Population Total	Source	Year	Exclusion From Control Totals
Australia[1]	16,704,354 (age 15-74)	Estimated resident population, projected from Census	2006	None
		Monthly Population Survey (MPS)	2011-2012	Members of the permanent defense forces, certain diplomatic personnel of overseas governments customarily excluded from census and estimated population counts, overseas residents in Australia, and members of non-Australian defense forces (and their dependents) stationed in Australia
		Survey of Education and Work (SEW)	2011	Ages 65-74, special dwelling type institutionalized persons, special dwelling type boarding school pupils, persons permanently unable to work, and persons living in collection districts that contain a discrete indigenous community in very remote areas
Austria	5,647,341	Population registry and Labor Force Survey	2011	Undocumented immigrants
Canada	23,381,067	Demographic projections of the Canadian population for April 2012 based on 2006 Census data	2012	Indian reserves in the provinces, institutions and non-institutional collective dwellings
Cyprus[7]	592,296	Census	2011	None
Czech Republic	7,395,111	Census	2011	Undocumented immigrants
Denmark	3,629,087	Registry	2011	Undocumented immigrants
England (UK)	34,257,191	Simple mean values for population estimates produced for each quarter in the calendar year 2011	2011	None
Estonia	896,163	Official Demographic Statistics	2012	Undocumented immigrants

[7] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 15-4 (cont.): Benchmark control totals, by country

Country	Population Total	Source	Year	Exclusion From Control Totals
Finland	3,496,909	Population database, education register for education level	2011	None
Flanders (Belgium)	4,138,042	Labor Force Survey	2010	None
France	4,0793,515	Labor Force Survey	2012	None
Germany	53,657,540	Microcensus	2010	Undocumented immigrants
Ireland	2,994,368	Census	2011	None
Italy	39,369,830	Italian Multipurpose Survey	2010	None
Japan	81,059,238	Census	2010	None
Korea	34,602,008	Census	2010	Undocumented immigrants, residents of small islands
Netherlands	11,160,541	Registry	2011, 2011-2012	Non-registered population
Northern Ireland (UK)	1,165,218	March 2010 population estimates	2010	None
Norway	3,282,755	Registry	2011	Undocumented immigrants
Poland	26,741,987	Registry	2011	Undocumented immigrants and foreigners staying in Poland fewer than 3 months
Russian Federation[8]	87,415,088	Census	2010	Moscow region and Moscow city
Slovak Republic	3,870,993	Census	2011	None
Spain	31,091,563	Registry	2012	None
Sweden	6,116,358	Registry	2011	Undocumented immigrants
United States	203,144,374	American Community Survey	2010	None

[1] Control totals were adjusted to meet the PIAAC scope, that is, all persons aged between 15 and 74 years old who do not live in very remote areas, special (i.e., nonprivate) dwellings, or collection districts that contain a discrete indigenous community, and exclude persons that are diplomatic personnel of overseas governments.

[8] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Deviations from weighting standards or excluded steps

The majority of countries performed each of the weighting steps described in section 15.1.2 and/or 15.1.3. The exceptions are enumerated in Table 15-5.

Table 15-5: Deviations from weighting standards or excluded steps, by country

Country	Weighting Deviations or Excluded Steps
Australia	Australia used person-level nonresponse adjustments and benchmarking to adjust for undercoverage and nonresponse at the household and person level, rather than performing a series of separate adjustments. Australia also applied an explicit trimming step, but if a weight was lower than 50% or higher than 300% of the initial weight after adjustments and benchmarking, benchmark classes were collapsed to reduce the weight fluctuation.
Austria	None
Canada	Canada's sample included several oversamples that were selected sequentially from the 2011 Canadian census or the 2011 National Household Survey databases, meaning that (1) there was an overlap between the frames used to select each sample, and (2) a unit selected for one part of the sample was no longer available for the other parts of the sample. As a result, the sum of weights of the whole sample would overestimate the size of the Canadian population aged between 16 and 65. Canada included an integration step at the end of the weighting process so that the final weights adequately represent the PIAAC population.
Cyprus[9]	None
Czech Republic	Weights for the Czech Republic main sample and supplemental sample were created separately and then composited at the end of the weighting process. In the supplemental sample, 30-year-olds were treated as 29-year-olds. The main, reserve and supplemental sample were selected in a sequential manner, and the screener base weights for the reserve and supplemental samples reflected conditional probabilities given the household was not selected for the previous sample. Therefore, the base weights for the sample main sample (including reserve) were adjusted downward so that they sum to the total of the base weights of the main sample without reserve. Following compositing, the weights for the combined samples were raked to ensure that the final composited weights agreed with the control totals used when raking the main sample.
Denmark	An unknown eligibility adjustment was not needed because Denmark did not have any inaccessible cases with unknown whereabouts.
England/N. Ireland (UK)	England/N. Ireland (UK) did not collect age and gender for all sampled persons during the screener. Therefore, in addition to the standard literacy-related nonresponse adjustment for screener countries, LRNRs with age and gender successfully collected represented those with age or gender not successfully collected. In addition, the theoretical person base weights (THEOR_PBWT) were derived from imputed values of the number of eligible people in the sampled household (NUM_ELG) for some cases due to a technical problem with the contact data that the interviewers entered.
Estonia	A literacy-related nonresponse adjustment was not needed for Estonia because all LRNRs had age and gender collected.
Finland	None

[9] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 15-5 (cont.): Deviations from weighting standards or excluded steps, by country

Country	Weighting Deviations or Excluded Steps
Flanders (Belgium)	An unknown eligibility adjustment was not needed because Flanders (Belgium) did not have any inaccessible cases with unknown whereabouts. A literacy-related nonresponse adjustment was not needed for Flanders (Belgium) because all LRNRs had age and gender collected.
Germany	Although the sample was probability based, Germany was unable to calculate exact selection probabilities due to an error in the sample selection algorithm. Therefore, the base weights were calculated using estimated probabilities from a simulation.
Ireland	None
Italy	None
Japan	A literacy-related nonresponse adjustment was not needed for Japan because all LRNRs had age and gender collected.
Korea	None
Netherlands	None
Norway	None
Poland	Poland did not collect age and gender for any of the BQ LRNRs and had very few assessment LRNRs, so the standard literacy-related nonresponse adjustment could not be performed. The BQ LRNRs together with the other BQ NRs were represented by BQ respondents. Poland's data were reweighted to correct for base weights. Poland discovered after weighting that in four cities the sample was not selected with equal probability (base weights adjusted to reflect differential selection probability) and a city was omitted during sample selection (base weights inflated for other cities with similar population to represent the omitted city). This led to more variability in their final weights.
Russian Federation[10]	A literacy-related nonresponse adjustment was not needed for the Russian Federation because there were no literacy-related nonrespondents at any stage of the data collection. Also, BQ nonresponse adjustment was not conducted because the BQ response rate was close to 100%.
Slovak Republic	An unknown eligibility adjustment was not needed because the Slovak Republic did not have any inaccessible cases with unknown whereabouts.
Spain	A literacy-related nonresponse adjustment was not needed for Spain because all LRNRs had age and gender collected.
Sweden	Sweden used benchmarking to adjust for undercoverage and nonresponse rather than performing a series of separate adjustments. To meet the requirements for the appropriate treatment of LRNRs, Sweden inflated the weights of assessment LRNRs to account for BQ LRNRs without age and gender collected. Then the base weights for the respondents were calibrated directly to known population totals (less the total for the LRNRs). Data collected from the survey (e.g., age) were not used in weighting, as all weighting variables were based on the registry data. After calibration, Sweden performed an unknown-eligibility adjustment to adjust for ineligibles since their population totals included ineligible cases.
United States	None

[10] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

15.2 Variance estimation

Inferences will not be valid unless the corresponding variance estimators appropriately reflect all of the complex features of the PIAAC sample design (e.g., stratification and clustering). The replication approach is used for estimating variances for the international analyses of PIAAC data. Under the replication approach, subsamples (also known as replicates) from the full sample are formed and statistics of the subsamples are used to estimate the variance of the full sample statistic. The replication approach, in conjunction with the multiple imputation approach used to derive the plausible values, captures the variation due to the complex sampling and estimation approaches, including:

- Sample design
- Selection
- Weighting adjustments
- Measurement error through the processing of multiple imputation of plausible values

For a detailed description on replication methods for different sample designs, refer to Appendix D of the WesVar® manual.[11]

The PIAAC Data Explorer is the primary tool for the analysis of PIAAC data. It has been adapted for handling the following four different replication schemes:

- Delete-one jackknife
- Paired jackknife
- Balanced repeated replication
- Fay's method

The delete-one jackknife is also referred to as delete-a-group jackknife, random groups approach or JK1. The paired jackknife is also referred to as JK2. The JK2 approach, with two variance units per stratum, is appropriate for sample designs where PSUs are stratified or selected with systematic sampling from a sorted list. The balanced repeated replication (BRR) approach is also commonly used when strata are involved, and Fay's method is a variant of the BRR approach.

Replication methods are applied to surveys by dividing the sample into specially designed replicate subsamples that mirror the design of the full sample. To form the replicate subsamples, variance strata and variance units are defined. Each subsample is reweighted to account for the subsampling that occurred. An estimate is then calculated for the full sample and each of the replicate subsamples. The variance of the full sample estimate is computed as the sum of squared deviations between each replicate subsample estimate and the full sample estimate. The general replication formula is

$$Var(\hat{\theta}) = c\sum_{i}(\hat{\theta}_i - \hat{\theta}_0)^2,$$

where

[11] http://www.westat.com/Westat/pdf/wesvar/WV_4-3_Manual.pdf

c	=	1,	for the paired jackknife (JK2)
	=	(g-1)/g,	for the random groups (delete-one) approach (JK1)
	=	1 / g	for the BRR approach
	=	$1/[g(1-k)^2]$	for Fay's method
g	=	number of replicates	
k	=	weighting factor for Fay's method	
$\hat{\theta}_0$	=	full sample estimate	
$\hat{\theta}_i$	=	estimate for replicate *i*.	

A variety of sample designs were employed across the different countries participating in PIAAC. Replication is adaptable to a wide variety of designs, including simple random sampling, systematic sampling, stratified designs and multistage cluster designs. In general, replication schemes are selected based on the sample design. A random groups approach may do well for a simple random sample while a paired jackknife mechanism is not meant for an SRS, but could be adapted. The paired jackknife would work very well for a one-PSU per stratum design, while a random groups design is not appropriate. Some efficiency is gained by selecting the most appropriate approach for the sample design.

15.2.1 Creation of replicate weights

Participating countries followed the PIAAC Technical Standards and Guidelines in providing the data necessary for creating replicate weights. All participating countries in PIAAC were responsible for defining variance strata and variance units. The specification of variance strata and variance units must conform to the design assumptions of a replication method and should be determined by the type of sampling design that was used to collect the data (e.g., whether or not stratification was used and how many PSUs were in each stratum). In addition, in some cases the sampling strata and PSUs had to be grouped to reduce the number of replicates to fit the sample design into a replication design that followed the PIAAC standards.

Once the variance strata and variance units were assigned, the Consortium/countries followed detailed guidelines on how to form and create the replicate weights. First, replicate base weights were created. For screener countries, the household base weights for the household were replicated. For registry countries, the person base weights were replicated. Subsequently, all weight adjustments that were conducted for the full sample were conducted on each replicate weight to capture the variation created, or reduced, by the weight adjustments.

15.2.2 Summary of country-specific variance estimation implementation

Table 15-6 presents the replication approach employed by each country. The choice of the replication method was guided by the particular sample design used in each country. For instance, JK1 is appropriate for a design that uses a registry without stratification or sorting. If strata were used and there were two primary sampling units (PSUs) per stratum, the appropriate replication method would be JK2, BRR or Fay's method. If there were many PSUs sampled from a small number of strata, then JK2, BRR or Fay's method could still have been used to reflect the sampling variation by creating pseudo-strata within the existing strata. The allowed number of replicates ranged from a minimum of 15 to a maximum of 80 replicate weights.

Table 15-6: Replication approach, by country

Country	First Stage Sample Design		Replication Method	Number of Replicates
	Stratification	Number of Sampled Units Per Stratum (for non-certainties)		
Australia	Yes	Not reported	JK1	60
Austria	Sorting only	NA	JK1	80
Canada	Yes	More than 2	JK1	80
Cyprus[12]	Yes	More than 2	JK2	80
Czech Republic	Yes	More than 2	JK2	80
Denmark[1]	Yes	More than 2	JK1	80
England (UK)	Yes	More than 2	JK2	80
Estonia	Yes	More than 2	JK2	80
Finland	Yes	More than 2	JK2	80
Flanders (Belgium)	Yes	More than 2	JK2	80
France	Yes	More than 2	JK2[2]	80
Germany	Yes[3]	0, 1, or 2	JK1	80
Ireland	Yes	More than 2	JK2	80
Italy	Yes	2	JK2	80
Japan	Yes	More than 2	JK2	80
Korea	Yes	More than 2	JK2	80
Netherlands	Sorting only	NA	JK2	80
Northern Ireland (UK)	Sorting only	NA	JK2	80
Norway	Yes	More than 2	JK2	80
Poland	Yes	More than 2	JK2	80
Russian Federation[13]	Yes	1, 2, 3, or 4	JK2	12[4]
Slovak Republic	Yes	More than 2	JK2	80
Spain	Yes	More than 2	JK2	80
Sweden	Yes	More than 2	JK2	80
United States	Yes	1	JK2	45

NA: not applicable; JK1: delete-one jackknife; JK2: paired jackknife.

[1] Denmark discovered an error in the calibration step after weighting had been completed (i.e., some population counts for the replicate calibration program were incorrect). The difference between the erroneous and the correct calibrated weights was less than 0.017 because the procedure calibrated to the correct population total. Because the impact on variances appeared to be small, no re-calibration was warranted.

[2] France's replicate weights were created using Fay's method. However, the variance computation can use the JK2 formula.

[3] Germany had a highly stratified design, with more strata than sampled PSUs.

[4] Due to the small number of PSUs selected, only 12 replicates could be formed for Russian Federation (11 from 22 noncertainty PSUs and 1 from 1 certainty PSU).

[12] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[13] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

15.2.3 Accounting for imputation error variance component

For estimation using plausible values (PVs), calculations must account for both the sampling error component and the variance due to imputation of proficiency scores. The estimator of the population mean is the average of the M PV means,

$$\hat{\bar{Y}}^{*} = \sum_{m=1}^{M} \hat{\bar{Y}}_m \Big/ M .$$

The variance of the estimated mean $\hat{\bar{Y}}^{*}$ is computed using formulas specific to PVs as follows:

$$v\left(\hat{\bar{Y}}^{*}\right) = U^{*} + B\left(1 + \frac{1}{M}\right)$$

where, the "within" variance component is computed as the average of the sampling variance for each of the M plausible values, computed as,

$$U^{*} = \left(\sum_{m=1}^{M} U_m\right) \Big/ M ,$$

where the sampling variance of the estimated mean $\hat{\bar{Y}}_m$ for plausible value m is U_m , and

where, the "between" component is calculated as

$$B = \left[\sum_{m=1}^{M} \left(\hat{\bar{Y}}_m - \hat{\bar{Y}}^{*}\right)^2\right] \Big/ (M-1) .$$

where, the mean of each of the M PVs $y_{l1}, y_{l2}, ..., y_{lM}$. for sample unit l is computed as

$$\hat{\bar{Y}}_m = \sum_{l \in s} w_l y_{lm} \Big/ \sum_{l \in s} w_l \; ; m = 1, ..., M ,$$

where s denotes the set of sample units.

The standard error is computed as the square root of the total variance, $\sqrt{v\left(\hat{\bar{Y}}^{*}\right)}$.

15.3 Recommendations for future cycles

Based on the Field Test and Main Study experience of PIAAC Round 1, the Consortium is proposing a series of recommendations for future cycles of PIAAC.

1. Countries should review the Weighting and Variance Estimation document during data collection and develop the programs needed for the completion of the Sample Design International File (SDIF).
2. More extensive quality checks should be conducted before countries submit the SDIF. The Consortium can provide such checks so they can be implemented by the country and/or incorporated into the Data Management Expert software.
3. Due to the complexities surrounding the assignment of the variance strata and variance units, for which the replicate weights are created, it is recommended that the Consortium conduct the assignment.

4. The Consortium will compute sample weights for all countries to ensure standardization unless a country has a reasonable justification (e.g., confidentiality issues) for weighting its own data.
5. Countries should review the set of variables used in weighting by other countries (Table 15-3) to see if any variables can be added to the weighting process for their country.
6. The same programs used for doing weight adjustments for the full sample weight must be used (or looped through) for each of the replicate weights. If the replicate weights are out of alignment with the full sample weights, it causes significant increase to the variances. This concern is dampened due to the recommendation that the Consortium conduct the weighting.
7. Countries need to ensure that the categories of the calibration variables, to be provided in the SDIF, are exactly the same (in terms of values and meaning) as given in the control totals.
8. Countries should conduct a comparison of control totals for two difference sources, explain the difference, and determine what is needed to be done for the control totals to have the same representation as the PIAAC target population.

References

Deming, W. E., & Stephan, F. F. (1940). On a least square adjustment of a sampled frequency table when the expected marginal totals are known. *Annals of Mathematical Statistics, 11,* pp. 427-444.

Lee, H. (1995). Outliers in business surveys. In B. Cox, D. Binder, B. Chinnappa, A. Christianson, M. Colledge & P. Kott (Eds.), *Business Survey Methods* (pp. 503-526). New York, NY: John Wiley & Sons.

Sarndal, C. E., Swenson, B., & Wretman, J. (1992). *Model Assisted Survey Sampling*. New York, NY: Springer-Verlag.

Chapter 16: Indicators of the Quality of the Sample Data

Leyla Mohadjer, Tom Krenzke and Wendy Van de Kerckhove, Westat

The sampling and weighting procedures described in Chapters 14 and 15 were undertaken with the goal of minimizing total survey error and producing samples that are representative of the target population. This chapter begins with a discussion of the quality assurance and quality control procedures that were implemented to ensure the sampling and weighting standards were met. The remaining sections report key quality indicators for each country. Section 16.2 provides coverage rates and response rates, section 16.3 describes the results of nonresponse bias analyses, and section 16.4 gives sample sizes and design effects.

16.1 Quality assurance and quality control procedures

Quality assurance (QA) and quality control (QC) procedures were put in place to ensure high-quality data that are comparable between countries. Section 16.1.1 describes the sampling-related QA process used by the Consortium to help achieve this goal. Section 16.1.2 describes the QC procedures required of countries to check that the quality goals related to sampling were met. Country compliance with the sampling, weighting, and nonresponse bias analysis QC procedures is addressed in sections 14.7 15.1.5, and 16.3, respectively.

16.1.1 Quality assurance activities

The QA process for sampling activities involved the development of standards and guidelines, production of sampling documents, creation of sampling and weighting activity toolkits, and communication with countries. This section provides a summary description of each activity.

Technical standards and guidelines for sampling and weighting

For Chapters 4 and 14 of the PIAAC Technical Standards and Guidelines, the Consortium produced standards, guidelines and recommendations for each of the following:

- **Target population:** To ensure that the target population for PIAAC is clearly defined in each country and is consistent across countries
- **Sampling frame:** To ensure that the sampling frame(s) is of high quality, provides acceptable coverage of the target population, and meets the requirements for sampling, location of selected population members, and estimation
- **Sample size:** To establish minimum sample size requirements for each country in order to meet the analysis goals of PIAAC
- **Sample design:** To specify the PIAAC sample design that will produce a probability-based sample, representative of the target population, in each participating country

- **Country-specific supplemental samples:** To describe potential country-specific supplemental sampling options and their implications for sample size
- **Sample selection:** To specify procedures for selecting a probability-based sample from the PIAAC target population following the sample design of PIAAC
- **Indicators of survey quality – noncoverage bias, nonresponse bias, and response rates:** To establish indicators to measure the quality of PIAAC survey data with respect to representation of the target population, and to provide standard procedures for measuring these indicators
- **Respondent incentives:** To increase response rates by offering sampled adults some incentive for participating in PIAAC and for attempting the assessment
- **Sample monitoring:** To monitor the sample during data collection, allowing timely reaction to any developing shortfalls or other potential for bias in the outcome sample
- **Weighting:** To provide a standard weighting approach and to facilitate the production of point estimates for the target population and their associated sampling error estimates

Sampling, weighting and nonresponse bias analysis (NRBA) documents

The Consortium created sampling, weighting and NRBA documents to provide further details on the quality standards in Chapters 4 and 14 of the PIAAC Technical Standards and Guidelines. The PIAAC Sampling Plan for the Field Test and PIAAC Sampling Plan (Main Survey) Part I gave an overview of the PIAAC sample design and a description of the information that countries should include in their sampling plan forms (described below). The PIAAC Weighting and Variance Estimation Plan described the weighting process, including the weighting steps, treatment of different disposition codes, calculation of weighting adjustment factors, assignment of variance strata and variance units, and creation of replicate weights. The paper entitled "PIAAC: Reducing Nonresponse Bias and Preliminary Nonresponse Bias Analysis" described the goals for identifying and reducing nonresponse bias before, during and after data collection. It also included requirements for the NRBA and examples of analyses conducted for past adult literacy surveys.

Sampling and Weighting Activity Toolkits

The Sampling and Weighting Activity Toolkits are a set of Consortium-developed programs and worksheets to aid countries in various sampling- and weighting-related activities. The toolkits were optional to countries but served to provide assistance to countries that needed it and helped ensure consistent and high quality results.

Types of toolkits included are as follows:

- Design effects (DEFF): Excel spreadsheets to compute DEFF due to clustering as well as DEFF due to differential sampling rates
- Within-household selection: Test input files for the algorithm to select one or two persons in a household
- Response rates: Excel spreadsheets to calculate actual and projected response rates for each data collection stage
- Variable selection: Programs, documentation, examples and test files for the selection of weighting variables

- Range of bias: Excel spreadsheet to evaluate the potential for nonresponse bias based on assumptions on how different nonrespondents are from respondents within the weighting classes

Sampling workshops and other communications

Communication with countries is an essential part of the QA process. To this end, the Consortium conducted a sampling workshop at the Barcelona (Spain) NPM meetings in March 2009. The workshop covered information on sample design, sampling plan forms, Field Test sampling requirements and sample sizes, Field Test quality control (QC) forms for sample selection and sample monitoring, and within-household selection. A second sampling workshop was held at the Princeton (NJ) NPM meeting in December 2010 that focused on lessons learned from the Field Test and preparing countries for the Main Study tasks of sample design and selection, weighting and variance estimation, and NRBA.

In February and March 2012, the Consortium held Web meetings to introduce the weighting QC forms (described below) and answer any weighting questions from countries. The sessions were offered at five different dates/times to accommodate country schedules. The Consortium also communicated with countries through presentations on sampling and survey operations requirements at NPM meetings and provided feedback through in-person consultation sessions (at NPM meetings) or through emails as needed.

16.1.2 Quality control activities

Sampling QC checks gathered information necessary to monitor the countries' sampling activities and facilitated a series of validity checks conducted by the Consortium. They were implemented through a series of electronic forms and data files for the Field Test and Main Study. The QC process started with the Consortium reviewing the materials and responding back to the country with suggestions for changes or recommendations for improvements. Each QC form or file had a submission schedule to ensure countries met the timeline for various project activities. Real-time monitoring of all aspects of sampling was critical in allowing the Consortium to uncover problems with sampling activities and for the countries to incorporate changes if necessary.

This section provides a summary description of each QC activity.

Sampling, weighting and NRBA plans

To reduce burden, the Consortium created a series of Sampling Plan Forms that contained all the information needed to meet the requirements listed in Chapter 4 (sample design and selection) and Chapter 14 (weighting/estimation) of the National Survey Design and Planning Report (NSDPR). Countries were required to complete and return the forms at least six months prior to the start of the Field Test data collection. This deadline was set to ensure Field Test sample design and selection steps provided all the necessary opportunities to test various aspects of the Main Study sample design and selection activities. Countries then had the opportunity to update their Main Study plans after the Field Test.

Sampling Plan Form Part 1 addressed the standards and guidelines related to sample design and selection. It was to be completed separately for the Field Test and Main Study. The form included questions on country plans for each of the following:

- Country-Specific Supplemental Samples
- Target Population Definition
- Background Design Information
- Sample Design and Sampling Units
- Within-Household Selection Rule (for countries with DU sampling)
- Sampling Frame Description
- Coverage Rate of Target Population
- Sample Selection Methods for Area Units (if applicable)
- Sample Selection Methods for DU and Within-Household Sampling (if applicable)
- Sample Selection Methods for Persons from Registries (if applicable)
- Sample Selection Checks
- Pre-Assignment of Assessment Instruments
- File Delivery
- Initial Sample Size Worksheet
- Reserve Sample
- Data Consistency Checks
- Sample Monitoring Plans
- Incentives

Sampling Plan Forms Part 2 and Part 3 pertained to the Main Study only. Part 2 checked countries' ability to comply with the weighting chapter (Chapter 14) of the PIAAC Technical Standards and Guidelines. It included questions on potential variables for weighting adjustments, planned weighting procedures, and the intended variance estimation method. Part 3 addressed expected response rates and NRBA plans.

Sample selection quality control forms

The QC sample selection (SS) forms collected detailed information about the country sample selection process and the results. Countries were to submit forms after each sample selection stage, allowing adequate time for countries to respond to the Consortium comments and questions and to revise procedures if necessary. The forms were important to verify that the selection of a probability sample adhered to the PIAAC Technical Standards and Guidelines.

The forms covered the following:

- Definition of the sampling unit
- Variables used for stratification, sorting and measure-of-size calculations
- List of certainty units, such as large primary sampling units
- Average, minimum and maximum cluster size
- Number of units on the frame, number of units sampled, weighted totals and target population totals, by stratum
- Weighted population totals by characteristics of interest (such as region or age)

- Weight distributions, where the weight is the inverse of the selection probability
- Description of any oversampling

Sample monitoring quality control forms

The sample monitoring process was intended to help countries identify potential shortfalls in the sample, problems in achieving the desired response rate, and the potential for nonresponse bias in the collected sample. Continuous monitoring was used to allow countries to employ procedures to address these problems during data collection while it was still possible to meet goals associated with sampling and data quality. Countries were required to complete QC sample monitoring (SM) forms every one to two months during data collection. The Consortium reviewed the forms and provided feedback to countries. The SM-1 forms collected information by key subgroups on the number of cases completed, response rates and expected yield. Countries were asked to monitor these figures by gender, age groups, geography and other characteristics of interest in order to help identify any shortfalls in yield or unusually low response rates. Starting mid-data collection, countries were also asked to provide a more extensive NRBA (SM-2) to identify subgroups with low response rates. The subgroups could be formed according to demographic or area-level characteristics believed to be related to proficiency. Multivariate techniques, such as a classification tree algorithm, were recommended for this evaluation to identify subgroups created from combinations of key variables.

Sampling-related quality control data checks

The Consortium provided countries with suggested sampling-related QC checks that the countries could run during data collection. These checks were intended to supplement the record consistency checks in the Data Management Expert (DME) software and emphasized variables relating to the Sample Design International File. Instructions were provided for checking consistency among disposition codes at the screener level (if applicable) and background questionnaire (BQ) level, checking the sampling of persons, and reviewing the conditions for a completed case as defined in standard 4.3.3.

Sample Design International File (SDIF) and Weighting International Files (WIFs)

At the end of data collection, countries provided the Consortium with an SDIF that contained sample selection data for each sampled unit, including sampling strata, probabilities of selection, ID variables, disposition codes, and auxiliary variables for weighting adjustments. The SDIF was the input file to the weighting process. The Consortium performed QC checks on the file to verify that variable definitions and formats were consistent with the specifications in Annex 4-3 of the PIAAC Technical Standards and Guidelines and that those fields reflected the information provided by the countries in their sample selection forms and weighting plans.

Countries also provided WIFs for Benchmark Control Totals to the Consortium. The files contained the external control totals to be used in the benchmarking adjustments. The benchmark WIFs were reviewed to check that the overall target population total was the same for each variable used in the benchmarking adjustment and that there was a set of control totals for each benchmarking variable included on the SDIF. Countries performing their own weighting adjustments also supplied a WIF for Quality Control Checks that was used to supplement the checks performed through the weighting QC forms (described below).

So as to not jeopardize the weighting schedule due to data reconciliation issues, countries were asked to provide a preliminary version of the SDIF and benchmark WIF before the end of data collection.

Weighting quality control forms

The Consortium developed a set of QC checks to review the weighting process for adherence to the weighting standards and guidelines and to check weight calculations for reasonableness and accuracy. Prior to the weighting period, each country needed to complete and return a checklist on the PIAAC Technical Standards and Guidelines related to weighting (Weighting QC Form W-0). They indicated whether the standards and guidelines were consistent with their implementation and understanding and indicated any deviations. They also needed to complete a W-1 form that contained checks on the base weights, variance strata and variance unit assignments, and any imputation performed for weighting variables.

Countries could opt to have the Consortium perform the weighting adjustments, or they could choose to create the final sampling weights themselves. During weighting, countries that formed their own weights were required to report on details of their weighting adjustments and weight distributions through a series of QC forms. If the Consortium conducted the weighting steps, the Consortium provided the forms to the countries for their review.

Form W-2 covered the household weights for countries with a household stage of sampling. Form W-3 was on the person-level weighting adjustments, and Form W-4 dealt with the final weights. The forms included the following checks:

- Descriptive statistics (including the counts of cases with missing and nonmissing weights, and sum, mean, minimum, maximum, and CV[1] of weights) on the full sample weights across weighting stages for all the sample, and by region, age group, and gender respectively
- Sum of replicate weights across weighting stages
- Descriptive statistics on selected replicate weights across weighting stages
- Unweighted and weighted counts by response status and weighting adjustment cells across weighting stages
- Description of trimming procedures
- Listing of the largest weights
- Comparison of control totals to external totals and weighted PIAAC totals
- Design effect calculations

Performing the weighting QC checks was essential for verifying that the final weights produced for estimation were appropriate. If any issues with the weighting adjustments were identified by the weighting QC forms, countries were required to rectify the problems and resubmit the QC forms until no more issues were found.

[1] Refer to section 16.4.2 for the definition of CV.

Weighted response rates and NRBA

Regardless of response rate, all countries were required to conduct a basic NRBA. The basic analysis evaluated the relationship of response status to available auxiliary variables and provides an indication of nonresponse bias prior to weighting adjustments. It could be used to inform the choice of weighting variables.

As described in section 16.2, the Consortium computed weighted responses rates for each country using the official response rate formulae in Annex 4-3 of the PIAAC Technical Standards and Guidelines and the data provided on the countries' SDIF. If a country's overall response rate fell below 70%, or if it had a stage of data collection with a response rate of less than 80%, the country was then asked to conduct an extended NRBA. This analysis included the evaluation of the potential for remaining bias after weighting adjustments were completed. It also attempted to evaluate bias directly in the proficiency estimates rather than solely relying on auxiliary variables.

Finally, countries were required to compute item response rates and conduct an item nonresponse bias analysis for any BQ items with response rates below 85%. The analyses were similar to those for the basic unit NRBA and involved comparing characteristics of item respondents and nonrespondents.

16.2 Sampling coverage and response rates

Coverage rates and response rates are important measures of the quality of the survey because they reflect the representation of the target population. Countries focused on reducing noncoverage and nonresponse bias given that the main goal of PIAAC is to produce high-quality unbiased estimates of the target population that are comparable across countries. First, section 16.2.1 contains an introduction to the implications of noncoverage and nonresponse on the potential for bias in the survey results. This will be discussed further in section 16.3. Then we turn to the computation of the coverage rates and the response rates.

16.2.1 Potential for bias

Under ideal situations, every eligible adult in the target population would have a nonzero chance of selection in a national sample, would be located and would agree to participate in the study. In practice, these circumstances are not realized in any survey population. There is a potential for bias whenever part of the target population is excluded from the frame or sampled persons who did not participate in the survey have different characteristics than those who did. For some important characteristics, the respondents may be substantially different from the rest of the target population, resulting in biased outcome estimates.

When response rates are low, there is a greater chance for nonresponse bias. The extent of nonresponse bias depends on how correlated the response propensity is with the survey outcomes. It is, therefore, critical to evaluate the potential for nonresponse bias, as a quality check on the estimates, at the conclusion of the data collection. Similarly, noncoverage bias (due to exclusions) can be substantial if the noncoverage rate is high and the difference in proficiency levels between adults included in the sample and those excluded from the frame is relatively large. Given the relationships between bias and coverage and response rates, countries had to keep the exclusion rates low and implement procedures to reduce the potential for nonresponse bias and attain high response rates.

The maximum allowable exclusion rate was set at 5% to guard against high noncoverage bias in PIAAC estimates. Any exclusions to the core PIAAC target population, whether or not they exceeded the threshold, were reviewed and approved by the Consortium. Even though up to 5% exclusions were tolerated, exclusions had to be kept to a minimum. If the quality of the sampling frame was such that it could result in a noncoverage rate of more than 5%, participating countries had to look into ways to improve coverage.

To reduce the potential for nonresponse bias, countries had to plan and implement field procedures that obtain a high level of cooperation. It was critical to monitor the distribution of the sample during data collection to ensure steps were taken to reduce the potential for bias as much as possible. As nonresponse rates increased, countries actively had to seek auxiliary data to reduce the impact of response propensities on the survey estimates. These auxiliary variables were used in weighting adjustments for the purpose of reducing nonresponse bias. Although sample weight adjustments based on auxiliary data are effective in reducing nonresponse bias, they are not considered as replacements for a vigorous effort to achieve the highest response rate possible.

16.2.2 Coverage rates

The PIAAC target population is defined as all noninstitutionalized adults between the ages of 16 and 65 (inclusive) who reside in the country at the time of data collection. The PIAAC Technical Standards and Guidelines require that the sampling frame covers at least 95% of the PIAAC target population. Exclusions (that is, persons who had no chance of being selected into the sample) may represent no more than 5% of the target population. There are, in effect, two categories of exclusions in PIAAC – *ex ante* exclusions by design (frame exclusions) and *ex post* exclusions following data collection (inaccessible persons). Both contribute to the overall noncoverage rate.

Exclusions by design

Exclusions by design or frame exclusions are of two types. They include, first, exclusions resulting from a decision not to include certain population groups in the sampling frame (e.g., the populations of remote and isolated regions) for reasons such as difficulty of access and the resulting high cost of data collection. Second, the use of a particular sampling frame may lead to the exclusion of certain groups in the population by virtue of the rules that determine which individuals are included in the list constituting the frame. For example, many population registers include only those members of the population with valid residence permits and, therefore, exclude illegal immigrants.

The frame noncoverage rate is computed as the estimated population in the excluded groups divided by the estimated core PIAAC target population. The rates by country are provided in Table 16-2. More information on sampling frame noncoverage, including the specific groups excluded by each country, is provided in Chapter 14.

Exclusions related to data collection

In addition to persons who are eligible under the international definition of PIAAC target population but were not included in the frame, persons that were included in the frame but in practice were impossible to be interviewed could be treated as exclusions. Some registry-based countries experienced difficulties locating and interviewing some or all sampled persons not residing at the address listed in the registry. Such cases were classified into a number of categories, as shown in Table 16-1. To arrive at an optimum and consistent approach across all registry-based countries, the Consortium assumed that all countries tried to find the location of the sampled persons and tried to interview them if they moved into one of the PSUs in the sample or were in a location where it was possible for PIAAC interviewers to visit and conduct the interview and assessment. Some individuals are found to be out of scope when the contact is attempted (e.g., information is provided that indicates that they have died, moved to an institutional setting, or emigrated). Others are “inaccessible” in that they cannot be interviewed because the information about their residential address was incorrect or because they have moved to another location in the country, which means they cannot be interviewed. Finally some members of the sample are untraceable in that no information about their whereabouts is available. The main advantage of classifying such cases in this manner was that the information about the inaccessible cases could be used to reduce the bias associated with noncoverage and, thus, reduce inconsistencies between country data.

The inaccessible noncoverage rate was calculated as the inaccessible population divided by the eligible population. The observed noncoverage rate had to incorporate sampling weights to account for selection probabilities and to ensure that the observed rate was representative of inaccessibles in the frame. If countries had an overall noncoverage rate (including frame and inaccessibles) of greater than 5%, up to 5% were reported in the noncoverage rate and the portion greater than 5% contributed as nonresponse in the response rate calculations.[2]

Table 16-2 shows the noncoverage rates for each country.

Table 16-1: Registry-based samples: Categories of ‘non-contacts’ and their status

Description	Status
Deceased	Out of scope
Moved outside country	Out of scope
Moved inside country	
Moved into institution	Out of scope
To PIAAC PSU	Inaccessible (unknown or invalid address)
To non-PIAAC PSU	Inaccessible (inability to interview outside PIAAC PSUs)
To unknown PSU	Inaccessible
Unknown whereabouts	Distributed between “out of scope” and “inaccessible” categories
Invalid address	Inaccessible

[2] This differs from the treatment of inaccessibles in weighting. For weighting purposes, such cases were treated as nonrespondents (see Chapter 15).

Table 16-2: Noncoverage rates: Sampling frame and inaccessible within sample

Country	Noncoverage Rate		
	Sampling Frame	Inaccessible	Overall
Australia	3.3%	0.0%	3.3%
Austria	0.6%	0.8%	1.4%
Canada	1.8%	0.0%	1.8%
Cyprus [3]	<2.0%	0.0%	<2.0%
Czech Republic	1.8%	0.0%	1.8%
Denmark	<0.1%	5.0%	5.0%
England (UK)	2.0%	0.0%	2.0%
Estonia	2.8%	0.6%	3.4%
Finland	0.2%	0.5%	0.7%
Flanders (Belgium)	1.0%	4.0%	5.0%
France	<2.6%	1.4%	<4.0%
Germany	0.5%	2.0%	2.5%
Ireland	0.4%	0.0%	0.4%
Italy	0.8%	1.9%	2.7%
Japan	2.2%	2.8%	5.0%
Korea	2.4%	0.0%	2.4%
Netherlands	0.9%	1.8%	2.7%
Northern Ireland (UK)	2.0%	0.0%	2.0%
Norway	0.4%	0.4%	0.8%
Poland	1.0%	4.0%	5.0%
Russian Federation[4]	1.5%	0.0%	1.5%
Slovak Republic	0.1%	4.9%	5.0%
Spain	0.0%	5.0%	5.0%
Sweden	<1.0%	0.0%	<1.0%
United States	0.1%	0.0%	0.1%

16.2.3 Response rates

Response rate is a valuable data quality measure and the most widely used indicator of survey quality. A high response rate increases the likelihood that the survey accurately represents the target population, and a low response rate reflects the possibility of bias in the outcome statistics.

A minimum overall response rate of 70% was set as the goal for PIAAC countries to be included in international indicators and reports, unless sample monitoring activities and/or nonresponse bias analyses indicate serious levels of bias in the country data. Countries with response rates of between 50% and 70% were included in international indicators and reports, unless other factors like noncoverage bias were detected. Deviations from the international standards on response rates were, however, documented in the international reports and publications. Results from countries with response rates below 50% were not published unless the country provided the

[3] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[4] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

OECD Secretariat with evidence that the potential bias introduced by the low response rates was unlikely to be greater than the bias associated with response rates of between 50% and 70%.

Using the standard formulae shown in Table 16-3, weighted response rates were computed hierarchically for the following stages of data collection:

- Screener (if the sample design included a screener stage)
- Background questionnaire
- Assessment (without and without reading components)
- Overall

Table 16-3: Response rate

Stage	Response Rate Calculation	Description
Screener	COMPLETE / ELIGIBLE COMPLETE = C^s ELIGIBLE = $HH^s - I^s - U^s * (I^s / K^s)$	C^s = Completed screeners, HH^s = All sampled households, I^s = HHs known to be ineligible, U^s = HHs with unknown eligibility status, K^s = HHs with known eligibility status.
Background Questionnaire (For countries with screeners)	COMPLETE / ELIGIBLE COMPLETE = $C^b + LR^b$ ELIGIBLE = $SP^b - D^b - I^b$	C^b = Completed BQ cases, LR^b = Literacy-related nonrespondents, SP^b = All sampled persons, D^b = SPs with a disability, I^b = SPs known to be ineligible.
Background Questionnaire (For countries with registries)	COMPLETE / (ELIGIBLE – EXCLUDE) COMPLETE = $C^b + LR^b$ ELIGIBLE = $SP^b - D^b - I^b - U^b * ((D^b + I^b)/ K^b)$ EXCLUDE = ELIGIBLE * EXC_PROP	C^b = Completed BQ cases, LR^b = Literacy-related nonrespondents, SP^b = All sampled persons, D^b = SPs with a disability, I^b = SPs known to be ineligible, U^b = SPs with unknown eligibility status, K^b = SPs with known eligibility status. EXC_PROP = Inaccessible rate from Table 16-2
Assessment[1]	COMPLETE / ELIGIBLE COMPLETE = $C^a + LR^a$ ELIGIBLE = $C^b - D^a - I^a$	C^a = Completed assessments, LR^a = Literacy-related nonrespondents, C^b = Completed BQ cases, D^a = SPs with a disability, I^a = SPs known to be ineligible.

[1] The assessment response rates with and without reading components were computed using the same formula, the difference being reflected in how each SP was classified, whether completing the reading components or not.

The literacy-related cases were included in the numerator of the response rates because their reason for nonresponse provides an indication of their proficiency level. The disabilities, while considered in scope, were subtracted from the denominator because the assessment did not accommodate such situations.

Table 16-4 shows a summary of the response rates for the participating countries.

Table 16-4: PIAAC response rates for participating countries

Country	Reading component	Response Rates					
		Without Reading Component				With reading component	
		Screener	BQ	Assessment	Overall	Assessment	Overall
Australia	Yes	85%	88%	96%	71%	96%	71%
Austria	Yes	-.-	53%	99%	53%	99%	53%
Canada[1]	Yes				59%		58%
Cyprus[5]	Yes	74%	99%	100%	73%	100%	73%
Czech Republic	Yes	74%	90%	100%	66%	100%	66%
Denmark	Yes	-.-	51%	97%	50%	97%	50%
England (UK)	Yes	89%	68%	97%	59%	97%	59%
Estonia	Yes	-.-	64%	99%	63%	99%	63%
Finland	No	-.-	69%	95%	66%	-.-	-.-
Flanders (Belgium)	Yes	-.-	62%	99%	62%	99%	62%
France	No	-.-	71%	94%	67%	-.-	-.-
Germany	Yes	-.-	55%	99%	55%	100%	55%
Ireland	Yes	79%	92%	99%	72%	99%	72%
Italy	Yes	88%	66%	97%	56%	97%	56%
Japan	No	-.-	50%	100%	50%	-.-	-.-
Korea	Yes	86%	91%	96%	75%	96%	75%
N. Ireland (UK)	Yes	83%	80%	98%	65%	98%	65%
Netherlands	Yes	-.-	53%	97%	51%	98%	51%
Norway	Yes	-.-	63%	98%	62%	98%	62%
Poland	Yes	-.-	56%	99%	56%	95%	54%
Russian Federation[6]	No	53%	99%	97%	52%	-.-	-.-
Slovak Republic	Yes	-.-	66%	99%	66%	99%	66%
Spain	Yes	-.-	48%	100%	48%	100%	48%
Sweden	Yes	-.-	46%	97%	45%	97%	45%
United States	Yes	86%	83%	99%	70%	99%	70%

[1] To account for multiple sampling frames and to provide an indication of nonresponse bias, nonresponse to the parent samples were reflected in Canada's PIAAC overall response rate computation. (See Chapter 14 for information on Canada's sample design.) It was decided that individual response rates at the screener, BQ and assessment stages are not to be reported.

[5] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[6] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Some countries chose to conduct the optional reading components as part of the psychometric assessment, and response rates with reading components were calculated for them. For all countries response rates were calculated without reading components, which provide a comparable measure across the countries. For countries with a screener, the overall response rate was calculated as the product of the response rates for the screener, BQ and assessment. For countries without a screener, the overall response rate was calculated as the product of the response rates for the BQ and the assessment. The screener response rate was weighted by the inverse of the household selection probability, and the BQ and assessment response rate by the inverse of the person selection probability. If countries had oversampling, it is reflected in the weights, and therefore weighted response rates are a comparable measure across countries.

16.3 Nonresponse bias analysis

Missing data can occur when some of the adults selected in the sample are not contacted or refuse to participate (referred to as unit nonresponse), they fail to respond to a particular survey item (referred to as item nonresponse), or because data collected from the sampled adults is contaminated (and thus not useful) or lost during or after the data collection phase. Nonresponse bias can be substantial when two conditions hold: 1) the response rate is relatively low and 2) the difference between the characteristics of respondents and those of nonrespondents is relatively large. This is reflected in the following deterministic nonresponse bias formula:

$$Bias(\bar{y}_R) = (1 - W_R)(\bar{Y}_R - \bar{Y}_{NR}),$$

where W_R is the proportion of respondents, $\bar{Y}_R$ is the mean outcome for respondents, and $\bar{Y}_{NR}$ is the mean outcome for nonrespondents. An alternative model of nonresponse assumes each sampled person has a certain propensity to respond, and nonresponse bias in a characteristic is a function of the covariance between the response propensity and the characteristic:

$$Bias(\bar{y}_R) = \frac{\sigma_{yp}}{\bar{p}},$$

where σ_{yp} is the covariance between the outcome variable and response propensity, and $\bar{p}$ is the mean response propensity. Based on this model, NRB is present if missingness is related to proficiency, as measured by PIAAC.

Countries worked to reduce nonresponse bias to the extent possible before, during, and after data collection. Before data collection, countries implemented field procedures with the goal of obtaining a high level of cooperation. Most countries followed the PIAAC required sample monitoring activities to reduce bias to the lowest level possible during data collection. Finally countries gathered and used auxiliary data to reduce bias in the outcome statistics through nonresponse adjustment weighting.

All countries were required to conduct a basic nonresponse bias analysis (NRBA) and report the results. The basic analysis was used to evaluate the potential for bias and to select variables for nonresponse adjustment weighting. In addition, countries were required to conduct and report the results of a more extensive NRBA if the overall response rate was below 70%, or if any stage of data collection (screener, background questionnaire, or the assessment) response rate was below 80%. An item NRBA was required for any BQ item with response rate below 85%.

A summary of the results of the basic NRBA is provided in Section 16.3.1. Section 16.3.2 contains the results of the extended NRBA, and Section 16.3.3 provides a summary of the item nonresponse analysis. A brief summary and conclusions of the NRBA is given in Section 16.3.4.

16.3.1 Basic NRBA

The basic NRBA involved comparing survey respondents and nonrespondents using auxiliary variables available on the sampling frame, available from a previous data collection stage (e.g. screener data for the BQ analysis), or coming from an external source that could be matched to each sampled unit. Also, observational data on respondents and nonrespondents collected during data collection could have been used to evaluate bias, assuming the data was of sufficient quality. The auxiliary variables must have been available for all eligible units and, as noted above, had to be related to proficiency. All countries were required to include the following variables in their analysis: age, gender, education, employment, and region. If any of these variables was not available for all eligible units, then a corresponding area-level variable could have been used instead (e.g. the employment rate within small geographic areas).

The basic analysis included results from the following:

- Comparison of response rates for different subgroups
- Use of a chi-square test or estimates of relative bias to compare the distribution of auxiliary variables (correlated with proficiency) for respondents and nonrespondents
- Use of a classification tree algorithm to identify subgroups with low response rates or use of logistic regression to model the relationship between response status and the auxiliary variables

The response rate and chi-square analyses were useful in explaining the relationship of response status to each auxiliary variable individually. A classification tree algorithm and/or a logistic regression model was used to evaluate the relationship between response status and multiple auxiliary variables.

All countries completed all the required analyses and included all the required variables, age, gender, education, employment, and region, in their analysis, with the exception of Austria, Finland, Flanders (Belgium) and Italy. In most cases, the failure to include the required variables in the analyses was due to the lack of access to sources with reliable data for such variables.

An initial basic NRBA was conducted prior to the weighting process. The analysis was conducted in two stages. The first stage helped to create a pool of predictor variables related to proficiency, using the field test data. The second stage helped to reduce the pool of predictor variables to those related to response propensity (this was repeated after the weighting process to finalize the basic NRBA). Most countries used all auxiliary variables that showed potential for bias in deriving nonresponse adjustments to the sampling weights. The remaining countries used most of the variables identified in the initial basic NRBA, mainly because reliable data was not available for the remaining variables.

Nonresponse weighting adjustments reduce bias in the outcome statistics to the extent that auxiliary variables are correlated with proficiency. Mainly, weighting adjustments are carried out by assuming nonrespondents' proficiency levels are the same as the respondents in the subgroups created for weighting adjustments using the auxiliary variables. This assumption is, of course, not true and the level of bias reduction depends on the number of auxiliary variables used during weighting and the correlation between these variables and proficiency.

The basic NRBA is a good initial assessment of nonresponse bias and is essential in identifying effective weighting variables. However, it has its limitations. The analysis does not reflect the effect of weighting adjustments on NRBA, and the extent of bias remaining after nonresponse adjustments are conducted. Therefore, countries with lower response rates were required to conduct a more extensive analysis to assess the potential for bias remaining after nonresponse adjustment weighting. Section 16.3.2 includes a brief description of the results of the extended NRBA.

16.3.2 Extended NRBA

A more extensive NRBA was required if the overall response rate was below 70%, or if any stage of data collection (screener, background questionnaire, or the assessment) response rate was below 80%.

Australia, Korea and the United States achieved an overall response rate of 70% or greater, with response rates for each stage being greater than 80%, and thus did not require the extended NRBA. Cyprus[7] and Ireland also achieved overall response rates of 70% or greater, but they achieved a lower than 80% response rate for one stage of their sample. The remaining countries achieved response rates lower than 70%.

The main purpose of the extended analysis was to assess potential for remaining bias in the final weighted proficiency estimates after adjusting for nonresponse. Because the proficiency levels of nonrespondents are unknown, the NRBA is carried out by making assumptions about nonrespondents. Therefore, it is necessary to conduct multiple analyses to assess the potential for bias since each analysis has its own limitations resulting from the specific assumptions made about nonrespondents. The extended NRBA included seven analyses (as listed below). Together, they were used to assess the patterns and potential for bias in each country data.

The extended NRBA included the following analysis:

1. Comparison of estimates before and after weighting adjustments;
2. Comparison of weighted estimates to external totals;
3. Correlations of auxiliary variables and proficiency estimates;
4. Comparison of estimates from alternative weighting adjustments;
5. Analysis of variables collected during data collection;

[7] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

6. Level-of-effort analysis; and
7. Calculation of the range of potential bias.

These analyses are described further below.

Cyprus[8] and Ireland were required to do only a subset of the analysis since their overall response rate was higher than 70%.

Comparison of estimates before and after weighting adjustments

To better capture the effects of the weighting adjustments on unit nonresponse bias, estimates from the full sample were compared to estimates from the respondents before and after weighting adjustments. To compare estimates before and after each step of weighting adjustments, the following comparisons were made:

- Comparison of percentage distributions from BQ base weights for the total eligible sample of persons with the BQ base weights for the BQ respondents to check for differences due to nonresponse to the BQ
- Comparison of percentage distributions from BQ base weights for the total eligible sample of persons with that from the BQ nonresponse adjusted weights for respondents to check for differences after the nonresponse adjustment process to the BQ
- Comparison of percentage distributions from BQ nonresponse adjusted weights for respondents with that from the BQ raked weights (weights adjusted to two or more marginal population totals) for respondents to check for differences that may have been introduced through the initial raking procedure

For countries that had screeners, analogous comparisons to the BQ level, as mentioned above, were completed. All the countries required to do the analysis completed it. The goal was to include at least one auxiliary variable not present in weighting adjustments in addition to those used during nonresponse adjustment weighting. Inclusion of the non-weighting variables shows whether the weighting adjustment was effective in reducing bias in other known auxiliary variables, not just the weighting variables. The following 11 countries; Denmark, England (UK), Finland, Germany, Japan, Netherlands, Northern Ireland (UK), Norway, Poland, Slovak Republic and Sweden and included nonweighting variables in this analysis as well as weighting variables. The remaining countries only included the weighting variables. Canada included a substantial number of weighting variables in their analysis. In general, all countries except for Russian Federation (partial compliance) observed that bias was reduced in auxiliary variables through weighting adjustments.

Comparison of weighted estimates to external totals

The second analysis compared estimates from PIAAC to external source estimates to assess potential for bias in PIAAC outcome statistics.

To the extent possible, countries used estimates from external sources that measured the same characteristic for a similar time period. Some external source estimates were subject to sampling

[8] See above footnote.

error also, and thus the variance of these estimates were taken into account when making comparisons.

Many countries found significant differences between the PIAAC estimates and the external source estimates but were mostly able to explain the sources for discrepancies. The sources mainly included, different data collection time periods or different definitions (e.g., definition of employment). All countries except France completed this analysis.

Correlations of auxiliary variables and proficiency estimates

The analyses described thus far relied on auxiliary variables and did not directly measure bias in the proficiency estimates. Bias in the auxiliary variables is indicative of bias in the proficiency estimates to the extent that the auxiliary variables and proficiency estimates are correlated. Thus, correlations between the auxiliary variables and proficiency data are good indicators of potential for bias reduction through weighting adjustments. For variables used in the weighting adjustments, a low correlation with proficiency implies that using the variable in the weighting adjustments did little to reduce nonresponse bias. On the other hand, a high correlation with proficiency implies a potentially high reduction in nonresponse bias. However, it should be noted that the correlations are based on respondents' data, and the relationship between proficiency and the auxiliary variables might be different for nonrespondents. Therefore, the correlations could be different if a country's response rate is very low, and if nonrespondents are different from respondents in terms of the relationship between their scores and the auxiliary variables.

Correlations were calculated as the square root of R-square of a weighted analysis of variance, whose dependent variable was the literacy or numeracy score while the explanatory variables were the weighting variables (BQ nonresponse adjustment cells and raking dimensions).

Table 16-5 presents the correlation between the proficiency and the weighting variables for each country.

Table 16-5: Correlations of auxiliary variables and proficiency estimates

Country	Literacy	Numeracy
Austria	0.56	0.57
Canada	0.54	0.53
Cyprus[9]	0.39	0.47
Czech Republic	0.56	0.60
Denmark	0.50	0.46
England (UK)*	0.52	0.56
Estonia	0.37	0.35
Finland	0.60	0.58
Flanders (Belgium)	0.36	0.36
France	0.60	0.64
Germany	0.61	0.62
Ireland	0.52	0.53
Italy	0.49	0.53

[9] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 16-5 (cont.): Correlations of auxiliary variables and proficiency estimates

Country	Literacy	Numeracy
Japan	0.53	0.52
Korea	0.55	0.55
Netherlands	0.57	0.55
Northern Ireland (UK)*	0.57	0.60
Norway**	0.48	0.48
Poland	0.40	0.37
Russian Federation[10]	0.35	0.34
Slovak Republic	0.38	0.38
Spain	0.62	0.62
Sweden	0.70	0.70
United States	0.63	0.66

*England (UK) and Northern Ireland (UK) were weighted separately to allow efficient estimates for each population.
** Norway was not able to provide nonresponse adjustment cells due to confidentiality concerns. Therefore, Norway self-reported the correlation between literacy scores and BQ nonresponse adjustment variables and raking variables as 0.48 for literacy. Norway did not report the correlation for numeracy. Therefore, 0.48 was assumed for numeracy.

There are a few countries with low correlation between the BQ nonresponse cells and the proficiency scores. However, all of the correlations between proficiency scores and the BQ nonresponse cells and the raking dimensions combined are higher than 0.30 and the average is 0.51 for literacy scores and 0.52 for numeracy scores. Although it was not required, the correlations for Korea and the US were also provided. Based on the moderate-to-high correlations between the weighting variables and the proficiency scores, we can expect the weighting adjustment to have reduced bias in the proficiency scores.

Figure 16-1 displays each country's correlation between weighting variables and the literacy score and correlation between weighting variables and the numeracy score. The two correlations are very close to each other, implying the same level of effectiveness in reducing bias for the two proficiency estimates.

[10] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Figure 16-1. Correlation of weighting variables and the proficiency scores

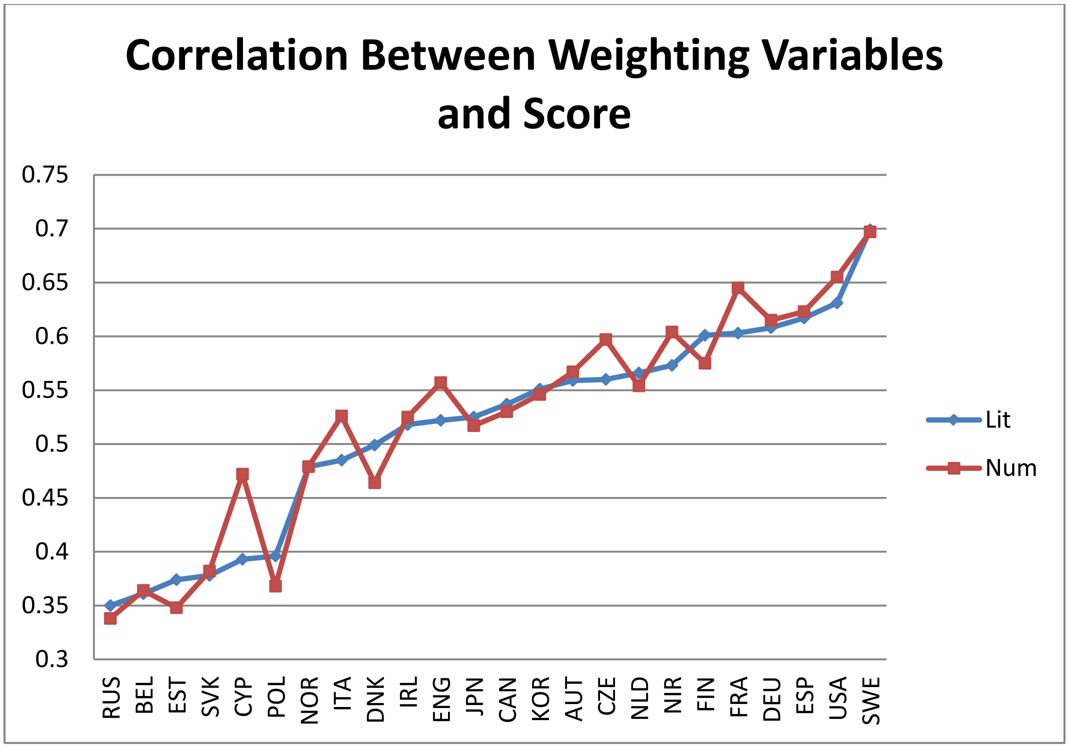


Figure 16-2 shows the plot of response rate versus correlation between the weighting variables and the literacy score reflecting the effectiveness of nonresponse adjustments in reducing bias.

Figure 16-2. Scatterplot of response rate versus correlation

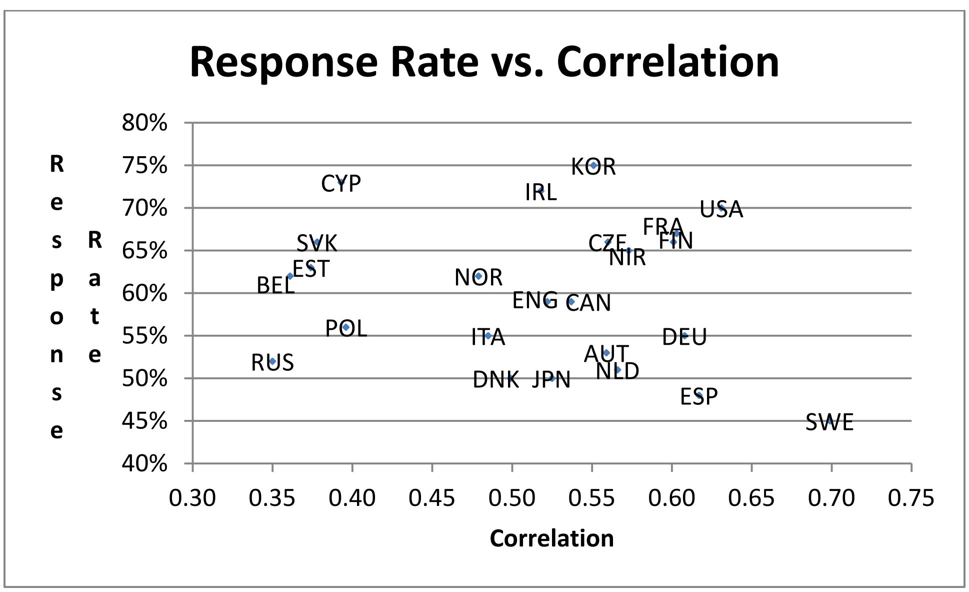


Figure 16-2 shows that:

- Countries in the lower right corner, such as Sweden, Spain and Germany, have low response rates, but are expected to have accomplished a considerable bias reduction through weighting, since their weighting variables are highly correlated with the proficiency.
- Austria, Canada, Denmark, England (UK), Italy, Japan and Netherlands have about average correlations, so bias reduction is expected at an average level as compared to other countries.
- Finland, France and the US have a higher than average correlation and high response rates.
- Cyprus,[11] Estonia, Flanders (Belgium) and Slovak Republic have low correlations, but relatively high response rates, which helped reduce potential for bias. Poland and Russian Federation, which also have low correlations, have somewhat lower response rates, which indicates relatively less potential for bias reduction.

Comparison of estimates from alternative weighting adjustments

For this evaluation, an auxiliary variable was re-calibrated to known totals, and estimates of the key statistics were compared before and after the re-weighting. Re-weighting was useful as an evaluation tool when:

- The variable was not used in weighting (because it was not available) or was used but with different categories
- The variable is correlated with the outcome measure
- The variable is correlated with response propensity

Any differences between estimates using the official survey weights and the re-weighted weights reflected noncoverage as well as nonresponse bias, but if there was not a large change in the estimates, this was further confirmation that nonresponse bias may not be a concern.

Thirteen of the countries fully complied with the analysis and results confirmed that nonresponse bias may not be a concern. These countries were: Austria, Canada, Denmark, Estonia, Finland, Flanders (Belgium), Germany, Japan, Netherlands, Norway, Poland, Spain and Sweden. Italy found a significant difference between the average literacy score using final weights and when using the alternative weights, where the alternative weights were created using a more detailed weighting variable. Some caution should be used in conclusions from this analysis for Czech Republic (quality unknown due to unavailability of data), France (did not comply), Russian Federation (did not comply), Slovak Republic (partial compliance) and UK (did not comply).

Japan and Sweden used the results of this analysis to improve their final survey weights.

Analysis of variables collected during data collection

Disposition codes contain information on reasons for nonresponse. For this analysis, distributions of sampled persons with known characteristics related to outcome (i.e. the literacy-related nonrespondent (LRNR) cases, which are language problems, reading and writing difficulty, and mental disability) were examined. For example, the demographic distribution of

[11] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

literacy-related cases was compared to other eligible persons using auxiliary data, and interview data. Statistical tests such as Chi-square tests were processed to determine if there is a relationship between select demographic variables and the disposition codes for nonrespondents. A special weighting adjustment for literacy-related cases was conducted for all countries, with the exception of Poland, where the BQ LRNRs together with the other BQ NRs were represented by BQ respondents. Therefore, in almost all countries, the existence of LRNR cases was dealt with appropriately in order to reduce potential for bias.

All countries, except for France, conducted an analysis of disposition codes with some observing differences that were expected, given the conditions in their countries. However, Sweden and the UK each conducted only a partially completed analysis (i.e., the quality level is unknown) due to unavailability of data.

In addition, Non-Interview Report (NIR) forms identify observable demographic information and reasons for nonresponse that are not captured in the disposition codes. The NIR forms can potentially indicate whether the reasons for nonresponse are related to proficiency estimates and suggest ways to improve response rates for future surveys.

The following countries put extra effort in conducting the analysis using the information from NIR forms: Cyprus[12], Germany, Italy, Japan, and Slovak Republic. The observed information from NIR forms may be useful for data collection in the next cycle.

Level-of-effort analysis

Another way to evaluate bias in the proficiency estimates is to compare proficiency estimates by level of effort. To the extent that the late or hard-to-reach respondents are similar to the nonrespondents, differences in proficiency estimates between the late and early (or hard-to-reach and easy-to-reach) respondents could indicate nonresponse bias. This analysis can be useful in detecting potential for bias given the assumption that nonrespondents are similar to respondents at the end of the data collection period.

If the literacy estimates differed between easy and hard respondents within a category of a weighting variable (used in the level-of-effort analysis), that may indicate that there are differences even within the weighting cells, and the nonresponse adjustment might not have helped. However, it may be that the data collection procedures were effective in obtaining a different type of respondent, potentially reducing the bias.

Thirteen countries revealed some significant differences in characteristics between early and late respondents, including Austria, Cyprus[13], Czech Republic, Denmark, Estonia, Flanders (Belgium), Italy, Japan, Netherlands, Norway, Poland, Spain and Sweden. Two countries, Finland and Germany, conducted the analysis but did not find significant differences. France, Russian Federation (due to the inability to classify respondents as difficult-to-contact) and Slovak Republic did not comply with the analysis, and some caution should be used in drawing conclusions from UK's analysis due to unavailability of data.

[12] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[13] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Calculation of the range of potential bias

The final component of the bias analysis is to evaluate the potential for bias remaining after weighting under the scenario that nonrespondents' proficiency scores are vastly different from the assumptions made during weighting.

It is well known that NRB can be reduced to some unknown extent through sample weighting when proficiency is correlated with auxiliary variables, and auxiliary variables are correlated with response propensity. Weighting assumes response probabilities are constant within every group created for weight adjustment, the proficiency score has zero variance within each group, and response propensity is uncorrelated with proficiency. It is known that these assumptions are not correct, and the impact of weight adjustments is limited to the number of variables available for nonresponse adjustment, and correlation levels with proficiency. Also, it is not possible to measure the exact departure from these assumptions since proficiency levels of nonrespondents are not known. This analysis attempts to evaluate the potential for bias by computing a range based on an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell. The range of bias was computed as the difference between the two extreme estimates, while taking into account the response rate and population size in the weighting cell.

The literacy scores' first plausible value was used to compute the range of scores within the responding sample and to predict the range of estimates for nonrespondents.

If the weighting classes were well defined, that is, each weighting class successfully contains a homogeneous population in terms of proficiency scores, then scores would not vary much within a weighting cell, so the range of bias would be small. On the other hand, the range of bias is also affected by the response rate. If the response rate is high, the range of bias may not be high even when the respondents have a wide range of scores in the weighting cell, because the proportion of nonrespondents whose score will get filled in with the extreme values is low. Thus, the range of bias analysis measures the impact of response rate on the quality of final estimates as well as the effectiveness of the weighting adjustments in reducing the potential for bias.

Figure 16-3 displays the range of potential bias in outcome statistics after weighting adjustments are incorporated in the official weights. For comparison purposes, the range of bias before weighting is included in the figure also. The range of bias before weighting was computed without regard to weighting cells, based on the extreme assumption that nonrespondents would all score at the 10th percentile, and at the other extreme they would all score at the 90th percentile. The countries are sorted by their nonresponse rate and each country's nonresponse rate is shown in Figure 16-4. Figure 16-3 shows that the range of bias after weighting adjustment is significantly lower than before weighting adjustments are conducted, that is, each country data achieved a substantial bias reduction through nonresponse adjustment weighting. In addition, countries with higher response rates, such as Ireland, France, Slovak Republic, and Estonia, have lower range of bias. However, some countries with a low response rate, such as Sweden, Spain, Denmark, and Japan, have low ranges of bias also, due to their effective nonresponse adjustment weighting processes. Results of from Russian Federation were inconclusive due to non-compliance. The results from the range of bias analysis re-emphasizes the importance of minimizing bias in the sample throughout the survey process, and achieving high response rates especially if the country does not have access to auxiliary variables highly correlated with proficiency.

Figure 16-3. Range of potential bias before and after weighting adjustment

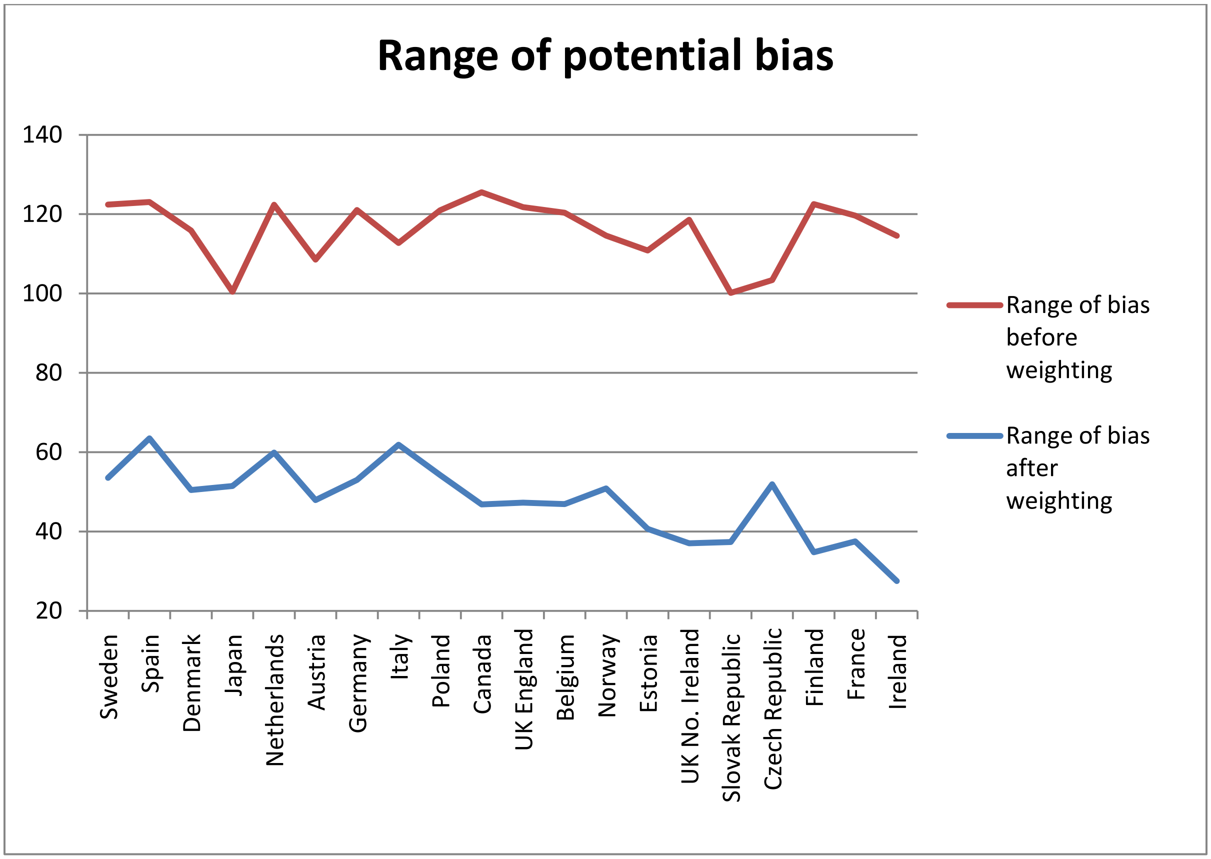
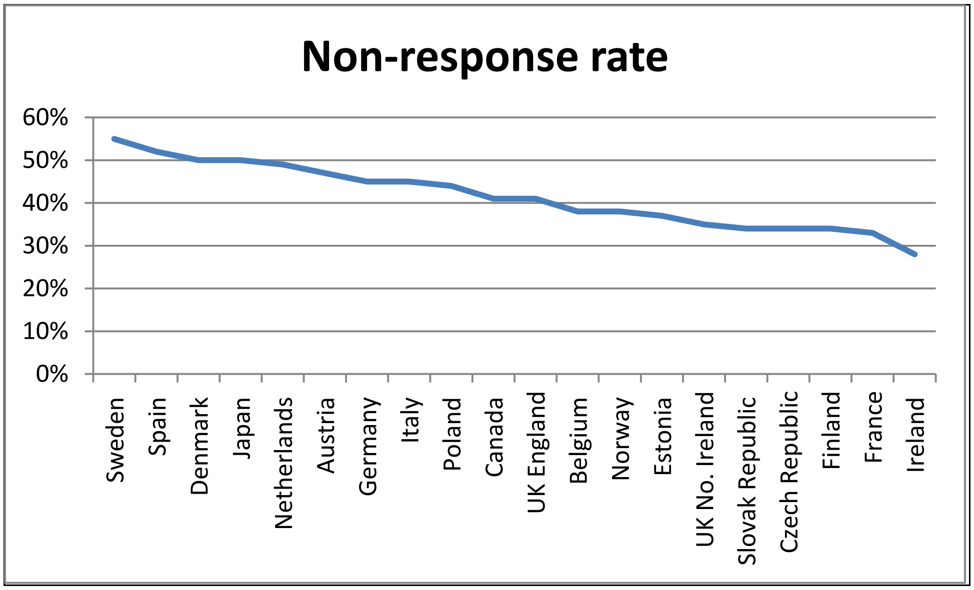


Figure 16-4. Nonresponse rate by participating country



16.3.3 Item NRBA

Countries were required to conduct nonresponse bias analysis for any BQ item with a response rate below 85%. Only two items showed low response rates: item D_Q17B (Earnings – additional payment amount last year), and item D_Q18A (Earnings – total earning last year).

Czech Republic, Estonia, Italy, Poland, Russian Federation and Slovak Republic were the only countries that had less than an 85% response rate for either D_Q17B or D_Q18A, with the lowest response rate being equal to 75% for D_Q18A for Poland.

16.3.4 Summary and conclusions

PIAAC standards were established with the main goal of producing reliable and comparable data across participating countries. As a result, a number of standards and guidelines were developed to help countries achieve the highest response rate possible, and at the same time reduce nonresponse bias to the minimum achievable. In addition, all countries were required to conduct a basic NRBA, and countries with lower response rates were required to conduct an extended NRBA.

All countries were required to conduct a basic nonresponse bias analysis (NRBA) and report the results. In addition, countries were required to conduct and report the results of a more extensive NRBA if the overall response rate was below 70%, or if any stage of data collection (screener, background questionnaire, or the assessment) response rate was below 80%. An item NRBA was required for any BQ item with response rate below 85%.

The basic and extended NRBA included several analyses. Each analysis was based a number of assumptions about nonrespondents, limiting the utility of the results. Thus, multiple analyses were used to assess the potential for bias in outcome statistics.

Correlation between the auxiliary variables used during weighting and the proficiency scores is a good indication of the effectiveness of nonresponse adjustment weighting. A number of countries with low response rates had higher correlations, implying a more effective nonresponse adjustment than countries with lower correlations. However, data users need to be cautioned that the analysis is based on correlations between respondents' proficiency scores and the auxiliary variables. That is, the analysis assumes that the same correlations exist for the remaining sampled cases that have no scores.

Table 16-6 summarizes the results of the NRBA for countries with response rates lower than 70%. The analysis showed that nonresponse adjustment weighting was effective in reducing the potential for bias in all countries. Countries that achieved higher response rates guaranteed a minimized level of bias in outcome statistics, whereas countries with lower response rates had to rely on the auxiliary variables available to them for nonresponse adjustment. Countries with relatively higher response rates and highly effective nonresponse adjustment showed minimal potential for bias as compared to countries with lower response rates, or those with moderately effective nonresponse adjustment weighting.

The analysis concluded that there was not enough evidence showing any moderate or high level of bias in the outcome statistics across the countries. However, this conclusion was based on assumptions made about the proficiency scores of nonrespondents. Therefore, data users need to be cautioned when interpreting the results of the NRBA for countries with very low response rates since different assumptions could lead to different results. For example, a response rate of 50% would mean making assumptions about half of the sample with no data. Multiple analyses, with different assumptions, were included in the NRBA to protect against misleading results, however, the lower the response rate, the higher is the risk of hidden biases that are undetectable through NRBA even when multiple analyses are involved.

Table 16-6: PIAAC NRBA outcome summary for countries with response rates lower than 70%

Country	Outcome
Austria	Caution-Bias low
Canada	Caution-Bias minimal
Czech Republic	Caution-Bias low
Denmark	Caution-Bias low
England (UK)	Caution-Bias low
Estonia	Caution-Bias low
Finland	Caution-Bias minimal
Flanders (Belgium)	Caution-Bias low
France	Caution-Bias minimal
Germany	Caution-Bias low
Italy	Caution-Bias low
Japan	Caution-Bias low
N. Ireland (UK)	Caution-Bias low
Netherlands	Caution-Bias low
Norway	Caution-Bias low
Poland	Caution-Bias low

Table 16-6 (cont.): PIAAC NRBA outcome summary for countries with response rates lower than 70%

Country	Outcome
Russian Federation [14]	Caution-Bias level unknown[1]
Slovak Republic	Caution-Bias low
Spain	Caution-Bias low
Sweden	Caution-Bias low

[1] Bias level unknown due to incomplete nonresponse bias analyses.

16.4 Sample sizes and design effects

A high-quality survey produces estimates that are both unbiased and low in variability. The bias aspect was discussed in previous sections. This section will address the variability aspect. Sample size is one of the main factors that affect the variability of survey estimates. The smaller the sample size, the higher the variability of survey estimates. However, given the same sample size, the survey estimates from a simple random sample often have lower variability than those from complex sample designs. The effect of the sampling design on the variability of estimates is usually referred to as the design effect. In the following, we discuss the PIAAC sample sizes and design effects in turn.

16.4.1 Sample sizes

Table 16-7 shows the actual sample size for each country. By "actual sample size", we refer to the number of cases with a final weight for analysis. The sample size includes both BQ respondents and BQ literacy-related nonrespondents (LRNR) with age and gender collected. The number of BQ LRNR cases is shown in a separate column as well. The BQ LRNR cases are different from the other nonrespondents because they did not complete the BQ due to literacy-related reasons, which means their proficiency levels cannot be represented by those of respondents. Therefore the percentage of such cases will be reported in data analysis although they do not have proficiency scores available.

Table 16-7: Actual sample sizes, by country

Country	Actual sample size*	BQ LRNR with age and gender collected
Australia	8,600	154
Austria	5,130	105
Canada	27,285	231
Cyprus[15]	5,053	661
Czech Republic	6,102	21
Denmark	7,328	42
England (UK)	5,131	51
Estonia	7,632	46
Finland	5,464	0
Flanders (Belgium)	5,463	480

[14] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.
[15] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 16-7 (cont.): Actual sample sizes, by country

Country	Actual sample size*	BQ LRNR with age and gender collected
France	6,993	86
Germany	5,465	86
Ireland	5,983	20
Italy	4,621	32
Japan	5,278	105
Korea	6,667	16
Netherlands	5,170	87
Northern Ireland (UK)	3,761	35
Norway	5,128	181
Poland	9,366	0
Russian Federation[16]	3,892	0
Slovak Republic	5,723	22
Spain	6,055	85
Sweden	4,469	0
United States	5,010	112

*The actual sample size is affected by several factors including response rates, number of languages, oversampling of subgroups, and the inclusion of reading components. Please refer to Chapter 14 for details.

16.4.2 Variability in sampling weights

A key component of the design effect is due to differential sampling weights. As mentioned in Chapter 14, several PIAAC countries sampled certain subgroups of population at a higher rate to obtain sufficient precision for analysis of the subgroups. For countries with a household sampling stage, people from different household sizes were also sampled with different probability. This led to unequal sampling weights and an increase in the variability of survey estimates. In addition, sampling weights were adjusted to account for sample nonresponse and undercoverage, which normally made the weights more variable. The variability of weights can be expressed by the coefficient of variation (CV) of the weights. The CV is

$$CV_w = \frac{\sigma_w}{\bar{w}},$$

where σ_w is the standard deviation of the weights and $\bar{w}$ is the mean of weights.

Table 16-8 shows the CV of both the base weights and final sampling weights for each country. The base weights are computed as the inverse of the probability of selection, while the final weights result from the weighting adjustments.

[16] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 16-8: Variability in sampling weights

Country	Sample Design	CV of household base weight[1]	CV of person base weight[2]	CV of person final weight
Australia	Screener	Not available[3]	Not available[3]	0.78
Austria	Registry	NA	0	0.30
Canada	Screener	1.31	1.28	1.33
Cyprus[17]	Screener	0.03	0.51	0.63
Czech Republic	Screener	1.52	1.71	1.37
Denmark	Registry	NA	0.46	0.52
England (UK)	Screener	0.30	0.57	0.59
Estonia	Registry	NA	0	0.21
Finland	Registry	NA	0.04	0.21
Flanders (Belgium)	Registry	NA	0	0.21
France	Registry	NA	0.10	0.23
Germany	Registry	NA	0.47	0.47
Ireland	Screener	0.37	0.62	0.61
Italy	Screener	0.12	0.50	0.66
Japan	Registry	NA	0.02	0.32
Korea	Screener	0.52	0.42	0.43
Netherlands	Registry	NA	0	0.31
Northern Ireland (UK)	Screener	0.82	2.29	0.73
Norway	Registry	NA	0	0.22
Poland	Registry	NA	0.91	0.97
Russian Federation[18]	Screener	0.57	1.44	1.04
Slovak Republic	Registry	NA	0	0.47
Spain	Registry	NA	0.33	0.46
Sweden	Registry	NA	0	0.36
United States	Screener	0.00	0.36	0.52

[1] Household base weights are not applicable (NA) to registry countries.
[2] For screener countries, the CV of person base weight is based on the person base weight described in section 15.1.3, which has the screener weighting adjustments in it.
[3] Australia did not provide information on the CVs of household and person base weight because of confidentiality restrictions.

The CV of the base weights is generally larger for countries with a household sampling stage (referred to as screener hereafter) than those without a household sampling stage (referred to as registry hereafter) due to differential probabilities of selection caused by differential household sizes. Among screener countries, the United Kingdom has the largest CV of base weights due to subsampling of multiple households at the same selected addresses in Northern Ireland (UK), and the Czech Republic's CV is high due to a supplemental sample of certain age groups. Among the registry countries, Poland has the largest CV caused by oversampling of certain age groups.

[17] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[18] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

16.4.3 Design effects and effective sample sizes

Many of the PIAAC countries used complex sample designs that involved clustered samples to meet cost limitations and be operationally feasible. For example, a sample may consist of 500 street blocks (clusters) with 10 people from each block. Because people who live in the same blocks tend to have more similar social and economic background than others, a simple random sample of 5,000 people is thus likely to cover the diversity of the population better than a sample of 500 blocks with 10 people from each block. Thus, the uncertainty (i.e. standard error) associated with any population parameter estimate will be larger for a clustered sample than for a simple random sample of the same size.

Furthermore, as mentioned in the previous section, unequal sampling weights also increased the variability of survey estimates.

The design effect is expressed by the ratio of the variance of the estimate obtained from the (more complex) sample to the variance of the estimate that would be obtained from a simple random sample with the same number of sampling units. Design effects can be used to evaluate the efficiency of the PIAAC sample designs. In addition, the design effects from this study can be used to estimate initial sample sizes for the next cycle of PIAAC.

As mentioned earlier in Chapter 15, the PIAAC variance can be estimated by using the replication technique[19], which accounts for the complex design (sampling and imputation error variance components as described in section 15), and a design effect can be computed for a statistic *t* using

$$Deff(t) = \frac{Var_{Complex}(t)}{Var_{SRS}(t)}$$

where $Var_{Complex}(t)$ is the variance for the complex sample for the statistic t computed by the replication method, and $Var_{SRS}(t)$ is the sampling variance for the same statistic on the same data but considering the sample as a simple random sample. The simple random sampling variance is computed as the average of the simple random sampling variance for each of the 10 plausible values.

Another way to express the reduction of precision due to the complex sample design is the effective sample size, which is the simple random sample size that would give the same sampling variance as the one obtained from the actual complex sample design. The effective sample size for a statistic *t* is

$$Effn(t) = \frac{n}{Deff(t)},$$

where *n* is the actual sample size.

The estimated design effects and effective sample sizes for proficiency scores over for each country are shown in Table 16-9 below.

[19] The Taylor Series linearization approach can be used to estimate the numerator as well.

Table 16-9: Design effects and effective sample sizes for proficiency score, by country

Country	Design effect			Effective sample size (Literacy)[1]
	Literacy	Numeracy	Problem solving	
Australia	2.39	2.06	2.81	3,061
Austria	1.41	1.61	1.44	3,561
Canada	3.45	4.39	4.80	7,848
Cyprus[20]	1.54	1.25	-.-	2,855
Czech Republic	3.53	2.75	2.87	1,725
Denmark	1.24	1.47	1.56	5,861
England (UK)	2.33	2.03	2.18	2,176
Estonia	2.00	1.02	2.95	3,785
Finland	0.94	1.00	1.73	5,464
Flanders (Belgium)	1.55	1.34	1.45	3,215
France	1.01	0.81	-.-	6,867
Germany	2.01	1.89	2.58	2,680
Ireland	2.25	2.16	2.57	2,652
Italy	2.75	2.08	-.-	1,666
Japan	1.54	1.48	2.38	3,362
Korea	1.31	1.52	2.02	5,086
Netherlands	1.10	0.99	1.50	4,635
Northern Ireland (UK)	6.62	4.71	7.14	563
Norway	0.83	1.05	0.88	4,947
Poland	1.48	2.47	4.54	6,320
Russian Federation[21]	15.77	16.62	22.33	247
Slovak Republic	1.35	1.58	1.74	4,236
Spain	1.27	0.88	-.-	4,710
Sweden	0.80	0.99	0.86	4,469
United States	2.21	2.05	2.84	2,211

[1] The effective sample size was computed as the number of cases with plausible values divided by the overall design effect. The effective sample size is set equal to the actual number of cases with plausible values for countries where the overall design effect is less than or equal to 1.

[20] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[21] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Chapter 17: Scaling PIAAC Cognitive Data

Kentaro Yamamoto, Lale Khorramdel and Matthias von Davier, ETS

17.1 Overview

The test design for PIAAC was based on a variant of matrix sampling (using different sets of items, multistage adaptive testing, and different assessment modes) where each respondent was administered a subset of items from the total item pool. That is, different groups of respondents answered different sets of items. That makes it inappropriate to use any statistic based on the number of correct responses in reporting the survey results. Differences in total scores (or statistics based on them) among respondents who took different sets of items may be due to variations in difficulty in the adaptively administered test forms. Unless one makes very strong assumptions – for example, that the different test forms are perfectly parallel – the performance of the two groups assessed in a matrix sampling arrangement cannot be directly compared using total-score statistics. Moreover, item-by-item reporting ignores the dissimilarities of proficiencies of subgroups to which the set of items was administered. Finally, using the average percentage of items answered correctly to estimate the mean proficiency of examinees in a given subpopulation does not provide any other information about the distribution of skills within that subpopulation (e.g., variances).

The limitations of conventional scoring methods can be overcome by using IRT scaling. When a set of items requires a given skill, the response patterns should show regularities that can be modeled using the underlying commonalities among the items. This regularity can be used to characterize respondents as well as items in terms of a common scale, even if not all respondents take identical sets of items. This makes it possible to describe distributions of performance in a population or subpopulation and to estimate the relationships between proficiency and background variables.

To increase the accuracy of the cognitive measurement, PIAAC uses plausible values (PVs) – which are multiple imputations – drawn from a posteriori distribution by combining the IRT scaling of the cognitive items with a latent regression model using information from the BQ (see chapters 3 and 20) in a population model.

In the following, the population model used for PIAAC scaling (IRT analysis, latent regression model, and computation of plausible values) is described formally (see section 17.2.). Its application to the PIAAC data is then demonstrated (see section 17.3.).

17.2 The latent regression item response model

This section reviews the scaling model employed in the analyses of the PIAAC data in theory – a latent regression item response model – and explains the multiple imputation or "plausible values" methodology that aims to increase the accuracy of the estimates of the proficiency distributions for various subpopulations and the population as a whole.

Most cognitive skills tests are concerned with accurately assessing the performance of individual respondents for the purposes of diagnosis, selection or placement. The accuracy of these measurements can be improved, meaning reducing the amount of measurement error, by increasing the number of items administered to the individual. Thus, achievement tests containing more than 70 items are common. Because the uncertainty associated with each estimated proficiency θ is negligible, the distribution of proficiency or the joint distribution of proficiency with other variables can be approximated using individual proficiencies. When analyzing the distribution of proficiencies for populations or subpopulations, however, more efficient estimates can be obtained from a matrix-sampling design.

In international large-scale assessments (ILSAs) such as PIAAC, test forms are kept relatively short to minimize individuals' response burden. At the same time, ILSAs aim to achieve broad coverage of the tested constructs. The full set of items is organized into different, but linked, assessment booklets; each individual receives only one booklet. Thus, the survey solicits relatively few responses from each respondent while maintaining a wide range of content representation when responses are aggregated. The advantage of estimating population characteristics more efficiently is offset by the inability to reliably measure and make precise statements about individuals' performance. Point estimates of proficiency that are (in some sense) optimal for each respondent could lead to seriously biased estimates of population characteristics (Wingersky, Kaplan, & Beaton, 1987). The "plausible value" methodology correctly accounts for error (or uncertainty) at the individual level by using multiple imputed proficiency values (plausible values) rather than assuming that this type of uncertainty is zero. Retaining this component of uncertainty requires that additional analysis procedures be used to estimate examinee proficiencies. This is done by applying a latent regression item response model to the data.

The latent regression item response model used for PIAAC incorporated test responses (responses to the cognitive items) as well as variables measured by the BQ (e.g., academic and nonacademic activities, and attitudes), which serve as covariates, in the computation of plausible values (von Davier, Sinharay, Oranje & Beaton, 2006). This approach was carried out as follows:

1) *Item calibration based on IRT*: An IRT model was fitted to the item responses. The responses consisted of dichotomous and polytomously scored values. These responses were used to calibrate the test and provide item parameter estimates for the (cognitive) test items.
2) *Population modeling using latent regressions and PV generation*: The population model assumes that item parameters are fixed at the values obtained in the calibration stage. Once the item parameters were estimated, a latent regression model was fitted to the data to obtain regression weights (Γ) and a residual variance-covariance matrix for the latent regression ($\sum$). Next, plausible values (Mislevy & Sheehan, 1987; von Davier, Gonzalez

& Mislevy, 2009) were obtained for all examinees using the item parameter estimates from the item calibration stage and the estimates of Γ and Σ from the latent regression model.

3) *Variance estimation*: To obtain a variance estimate for the proficiency means of each country and other statistics of interest, a replication approach (see, e.g. Johnson, 1989; Johnson & Rust, 1992) was used to estimate the sampling variability as well as the imputation variance associated with the plausible values.

The analytic procedures that establish these three modeling stages are explained further in the following sections.

17.2.1 Item response theory (item calibration)

PIAAC used the two-parameter logistic model (2PL; Birnbaum, 1968) for dichotomously scored responses and the generalized partial credit model (GPCM; Muraki, 1992) for items with more than two response categories.

The *2PL model* is a mathematical model for the probability that an individual will respond correctly to a particular item from a single domain of items. The probability of solving an item depends only on the respondent's ability or proficiency and two item parameters characterizing the properties of the item (item difficulty and item discrimination).The probability is given as a function of this person parameter and the two item parameters; it can be written as follows:

$$P\left(x_{ij} = 1 \middle| \theta_j, \beta_i, \alpha_i\right) = \frac{\exp\left(\alpha_i\left(\theta_j - \beta_i\right)\right)}{1 + \exp\left(\alpha_i\left(\theta_j - \beta_i\right)\right)}$$

where

x_{ij} is the response of person j to item i, 1 if correct and 0 if incorrect;

θ_j is the proficiency of person j (note that a person with higher proficiency has a greater probability of responding correctly);

α_i is the slope parameter of item i, characterizing its sensitivity to proficiency (item discrimination);

β_i is its locator parameter, characterizing item difficulty.

Note that, for $\alpha_i > 0.0$ this is a monotone increasing function with respect to θ; that is, the conditional probability of a correct response increases as the value of θ increases. In addition, a linear indeterminacy exists with respect to the values of θ_j, α_i, and β_i for a scale defined under the 2PL model. In other words, for an arbitrary linear transformation of θ say $\theta^* = A\theta + B$, the corresponding transformations $\alpha^*_i = \alpha_i / A$ and $\beta^*_i = A\beta_i + B$ give:

$$P\left(x_{ij} = 1 \middle| \theta_j^*, \beta_i^*, \alpha_i^*\right) = P\left(x_{ij} = 1 \middle| \theta_j, \beta_i, \alpha_i\right)$$

A central assumption of IRT is conditional independence (sometimes also called local independence). In other words, item response probabilities depend only on θ and the specified item parameters – there is no dependence on any demographic characteristics of the examinees, or responses to any other items presented in a test, or the survey administration conditions.

Moreover, the 2PL model assumes unidimensionality, that is, a single latent variable, θ, accounts for performance on a set of items. This enables the formulation of the following joint probability of a particular response pattern $x = (x_1, \ldots, x_n)$ across a set of n items.

$$P(x|\theta,\beta,\alpha) = \prod_{i=1}^{n} P_i(\theta)^{x_i}\,(1 - P_i(\theta))^{1-x_i}$$

When replacing the hypothetical response pattern with the scored observed data, the above function can be viewed as a likelihood function that is to be maximized with respect to the item parameters. To do this, it is assumed that respondents provide their answers independently of one another and that the respondent's proficiencies are sampled from a distribution $f(\theta)$. The likelihood function is characterized as

$$P(x|\boldsymbol{\beta},\boldsymbol{\alpha}) = \prod_{j=1}^{J}\int\left(\prod_{i=1}^{n} P_i(\theta_j)^{x_{ij}}\,(1 - P_i(\theta_j))^{1-x_{ij}}\right) f(\theta)d\theta$$

The item parameters obtained by maximizing this function are used in the subsequent analyses.

The GPCM (Muraki, 1992), like the 2PL, is a mathematical model for the probability that an individual will respond in a certain response category on a particular item. While the 2PL is suitable for dichotomous responses only, the GPCM can be used with polytomous and dichotomous responses. The GPCM reduces to the 2PL when applied to dichotomous responses. For an item *i* with *m*+1 categories, the model equation of the GPCM can be written as:

$$P(x_i = s|\theta_j, \beta_i, \alpha_i) = \frac{\exp\,(s\alpha_i\theta_j - \sum_{r=1}^{s}\beta_{ir})}{1+\sum_{u=1}^{m}\exp\,(s\alpha_i\theta_j - \sum_{r=1}^{u}\beta_{ir})}$$

Although the assumption of unidimensionality for the 2PL and GPCM may be considered a strong assumption, the use of these models is motivated by the need to summarize overall performance parsimoniously within a single domain. Hence, item parameters are estimated for each skill scale separately.

A critical part of the data analysis involves testing the assumptions of the 2PL, especially the assumption of conditional independence and the assumption of unidimensionality. Conditional independence means that respondents at a given ability level have the same probability of producing a correct response on an item regardless of their responses to other items as well as other attributes, including background variables such as citizenship, gender, immigrant status. Serious violation of the conditional independence assumption would undermine the accuracy and integrity of the results.

It is not uncommon for some items to violate this assumption. One expression of these types of model violations is differential item functioning (DIF), which means that items are either unsuitable, or much harder or easier, for a particular subpopulation compared to the other groups within the population. While the item parameters were being estimated, empirical conditional percentage-correct statistics were monitored across the samples to test for DIF in PIAAC. More precisely, for each item, the empirical item characteristic curves (ICC) for each country were compared to the expected ICC of the item. If the empirical ICCs for a certain item differed noticeable from the expected ICC, this would be evidence of DIF. For each country, a few items were identified that showed DIF in the international calibration (see section 17.3.2) and thus, did not conform to the common (international) item parameters.

Country-specific item parameters (computing national calibrations; see section 17.3.2) for items exhibiting country-level DIF in the international calibration were estimated to reduce potential bias introduced by these deviations. This approach was favored over dropping the country-specific item responses for these items from the analysis in order to retain the information from these responses. While the items with country DIF treated in this way no longer contribute to the international set of comparable responses, they continue to contribute to the reduction of measurement uncertainty for the specific country.

The software used for calibration, *mdltm* (von Davier, 2005), was enhanced by implementation of an algorithm that monitored DIF measures and that automatically generated a suggested list of country specific item treatments. This algorithm grouped similar deviations of subgroups of countries so that unique parameters were assigned to either individual countries or country groups that showed the same level and direction of deviation.

17.2.2 Population modeling using latent regressions and plausible values

The population model used for PIAAC is a combination of an IRT model and a latent regression model. Following item calibration using the IRT model, a latent regression model was estimated in a second step to obtain regression weights (Γ) and the residual variance-covariance matrix for the latent regression ($\sum$).

The latent regression model of Θ on Y with $\Gamma = (\gamma_{ij}, i = 1,\ldots,k; j = 0,\ldots,L)$, $Y = (1, y_1, \ldots, y_L)^t$, and $\Theta = (\theta_1, \ldots, \theta_K)^t$ can be described as follows:

$$\theta_i = \gamma_{i0} + \gamma_{i1} y_1 + \ldots + \gamma_{iL} y_L + \varepsilon_i,\ i = 1, \ldots, K$$

where ε_i is an error term.

The residual variance-covariance matrix can then be described with the following equation:

$$\Sigma = \Theta\Theta^t - \Gamma(YY^t)\Gamma^t$$

Both the item parameters from the calibration stage and the estimates from the regression analysis are needed to generate plausible values.

In the latent regression model, the distribution of the proficiency variable (θ) is assumed to depend not only on the cognitive item responses (x) but also on a number of predictors (y), which are variables obtained from the BQ (e.g., gender, country of birth, education, occupation, employment status, reading practices, etc.).

Usually, a considerable number of background variables (predictors) are collected in ILSAs, with a principal component analysis extracting the components that explain 90% of the variation for further analysis. In PIAAC it was decided to use 80% of explained variance to avoid overparameterization; (see section 17.3.4.). The use of principal components also serves to retain information for examinees with missing responses to one or more background variables. For the regression of the background variables on the proficiency variable it is assumed that:

$$\theta \sim N(y\Gamma, \Sigma)$$

The latent regression parameters Γ and Σ are estimated conditional on the previously determined item parameter estimates (from the item calibration stage).

Plausible values for each respondent *j* are drawn from the conditional distribution

$$P(\theta_j | x_j, y_j, \Gamma, \Sigma)$$

where Γ is the matrix of regression coefficients and Σ is a common residual variance matrix. Using standard rules of probability, the conditional probability of proficiency can be represented as follows:

$$P(\overline{\theta}_j | x_j, y_j, \Gamma, \Sigma) \propto P(x_j | \overline{\theta}_j, y_j, \Gamma, \Sigma) P(\overline{\theta}_j | y_j, \Gamma, \Sigma) = P(x_j | \overline{\theta}_j) P(\overline{\theta}_j | y_j, \Gamma, \Sigma)$$

where θ_j is a vector of scale values (these values correspond to performance on each of the three skills), $P(x_j|\theta_j)$ is the product over the scales of the independent likelihoods induced by responses to items within each scale, and $P(\theta_j|y_j, \Gamma, \Sigma)$ is the multivariate joint density of proficiencies of the scales, conditional on the observed value y_j of background responses and parameters Γ and Σ. The item parameters are fixed and regarded as population values in the computation described in this section.

The basic method for estimating Γ and Σ using the expectation-maximization (EM) algorithm is described in Mislevy (1985) for the single scale case. The EM algorithm requires the computation of the mean and variance, of the posterior distribution in (10).

After the estimation of Γ and Σ is complete, plausible values are drawn in a three-step process from the joint distribution of the values of Γ for all sampled respondents. First, a value of Γ is drawn from a normal approximation to $P(\Gamma,\Sigma|x_j,y_j)$ that fixes Σ at the value $\hat{\Sigma}$ (Thomas, 1993). Second, conditional on the generated value of Γ (and the fixed value of $\Sigma = \hat{\Sigma}$), the mean m_j^p, and variance Σ_j^p of the posterior distribution are computed using the same methods applied in the EM algorithm. In the third step, the θ are drawn independently from a multivariate normal distribution with mean m_j^p and variance Σ_j^p. These three steps were repeated 10 times, producing 10 imputations of θ for each sampled respondent (see section 17.3.4.).

The software DGROUP (Rogers et al., 2010) was used to estimate the latent regression model and generate plausible values. A multidimensional variant of the latent regression model was used that is based on Laplace approximation (Thomas, 1993).

17.3 Application to PIAAC

This section illustrates an application of the different steps of the population modeling described above using the PIAAC Main Study data. First, an overview of the data preparation is given. Then the national and international item calibration using the 2PL and the GPCM is described, as well as the computation of plausible values and their transformation onto the reporting scale. More specifically, the procedures utilized for the linking, with the aim to obtain equivalent scales, are described.

Scaling and analyses of the PIAAC data were carried out separately for each of the domains literacy, numeracy, and problem solving in technology-rich environments. By creating a separate scale for each, it remains possible to explore potential differences in subpopulation performance across these skills.

17.3.1 Sample size, data preparation, scoring, handling of missing values, block order effects

The following section provides an overview of the sample size, the number of items in the PIAAC assessment, the scoring and handling of missing values, and the examination of block order effects.

<u>Sample size</u>

PIAAC collected competency (cognitive) information through a series of assessment booklets containing literacy, numeracy and problem-solving tasks, and descriptive information through a BQ. Respondents were sampled using a stratified sampling method. Each participating country received instructions for sampling, weighting and data collection. However, each country carried out the actual design and administration of data collection activities separately.

PIAAC respondents' ages ranged from 16 to 65. Eligible participants included individuals who were living in households; institutional populations were excluded. Australia included participants younger than 16 and older than 65 in its target population, but these respondents were excluded from the PIAAC scaling process. Thus, tables comparing proficiency distributions of countries only include respondents between the ages of 16 and 65.

As with ALL, most countries used a modest monetary incentive in PIAAC. Without incentives, the participation rate may have been low enough to undermine the comparability of results.

Twenty-four countries participated in PIAAC (see Table 17.1). All 24 countries were asked to deliver their data before a certain deadline in order to allow sufficient time for analysis and reporting. Data from 331,863 respondents were received; the weighted data from 165,599 respondents between the age of 16 and 65 were available for statistical analyses (after data cleaning).

Table 17.1: Participating countries in PIAAC and sample sizes

Country	Sample Size (n)	Country	Sample Size (n)
Australia	7,430	Italy	4,621
Austria	5,130	Japan	5,278
Canada	27,285	Korea, Republic of	6,667
Canada (English)	21,374	Netherlands	5,170
Canade (French)	5,911	Norway	5,128
Cyprus[1]	5,053	Poland	9,366
Czech Republic	6,102	Russian Federation[2]	3,892
Denmark	7,328	Slovak Republic	5,723
Estonia	7,632	Spain	6,055
Finland	5,464	Sweden	4,469
Flanders (Belgium)	5,463	United Kingdom	8,892
France	6,993	*England (UK)*	5,131
Germany	5,465	*N. Ireland (UK)*	3,761
Ireland	5,983	United States of America	5,010

Assessment mode, testing time, item number and response format:

PIAAC was composed of a BQ and a core set of questions focusing on ICT applied through an interview using a computer-assisted format, and a cognitive assessment measuring the three domains. Based on the information from the BQ, the cognitive assessment was administered with either a CBA or PBA. Table 17.2 provides an overview of the frequency of selection and routing of respondents into these assessment modes.

Table 17.2: Proportion of the application of the assessment modes by domain in PIAAC

Domain	PBA (%)	CBA (%)	PBA+CBA (%)
Core	**22.8**	**73.9**	**96.7**
Literacy	**10.6**	**50.8**	**61.4**
Numeracy	**10.4**	**50.9**	**61.3**
Problem Solving	**NA**	**33.7**	**33.7**

[1] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[2] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

The BQ consisted of 258 variables (measured by more than 258 items, often exceeding 400 items; different countries had a different number of BQ items due to different country specific needs) measuring demographic characteristics, educational experiences, labor market experiences, and activities related to the assessed skills. In general, these questions did not require respondents to read any materials; they were administered by an interviewer, and only those questions that are applicable to the respondents' background were presented (see also chapters 3 and 20). Thus a respondent's reading proficiency was not a primary factor in the collection of the background information. In cases where the selected respondent was unable to speak the official language, another household member was permitted to act as an interpreter between interviewer and respondent for the collection of the background information only. Responses to the background questions served two major purposes. First, they provide a way to summarize the survey results using an array of descriptive variables, such as gender, age, educational attainment and country of birth. Second, they were used in the population model to increase the accuracy of the proficiency estimates for various subpopulations as described in section 17.2.

The *ICT core and the domain-based core part* are described in more detail in Chapter 1 of this volume. These sets of core items were used in selecting the paper or computer path for the respondents as well as the level of the computer-based stages in the subsequent assessment.

The *cognitive assessment* consisted of 166 items: literacy (76 items), numeracy (76 items), and problem solving (14 items). An additional 100 items measuring reading component skills were administered in a PBA if respondents failed to succeed in the other cognitive domains, for a total of 266 items in the cognitive assessment pool. Table 17.3 provides an overview of the number of items per cognitive domain and assessment mode. The large number of items was necessary to achieve adequate content coverage for each domain.

Table 17.3: Number of cognitive items per assessment mode and domain in PIAAC

Domain (Subscale)	Assessment Mode	Number of Items
Literacy	CBA	52
	PBA	24
Numeracy	CBA	52
	PBA	24
Problem Solving	CBA	14
Reading Components	PBA	100

Note: 18 literacy and 17 numeracy items were linking items between the PBA and CBA assessment mode, meaning these items were identical; thus PIAAC contained a total of 131 unique items

Each individual assessment started with the BQ, followed by the core items, and finished with the cognitive assessment. Each survey participant spent approximately 75-100 minutes on the entire assessment:

- BQ and ICT core items: 25-40 minutes

- cognitive assessment (including core items), one booklet: 50-60 minutes (341 booklets: four paper-based booklets and 337 computer-based booklets/paths; see Chapter 1)

The cognitive items were administered using either short open-ended response formats on paper or computer-based open response formats (e.g. highlighting the correct phrase or word); responses were classified into four categories: correct, incorrect, omitted, and not presented.

Scoring and handling of missing data

The 76 literacy items, 76 numeracy items, and 100 reading component items were dichotomously scored (solved: 1, not solved: 0), while the 14 problem-solving items were dichotomously or polytomously scored (five 3-point, one 2-point, and eight dichotomously scored items). For the problem-solving items, an automated scoring algorithm was used to score the responses from the CBA. One of the innovations introduced in PIAAC was the use of the LCS algorithm (longest common subsequence); this algorithm allowed for a scoring method that is automated yet emulates the leniency shown by human scorers in cases where underlining or highlighting responses would typically be evaluated. Humans recognize with ease if a respondent highlights or underlines the correct phrase even if they carelessly error omit one or two characters at the end of the line, at the beginning, or somewhere in the middle of the text. The LCS was used in conjunction with a discrepancy measure to allow for scoring of these “almost complete” responses in a comparable way across countries. As part of this process, a country-and-language independent threshold was established for each item based on the rationale that reasonably small deviations from the completely correct underlining should be considered as correct responses (Sukkarieh, von Davier & Yamamoto, 2012).

Regarding the handling of missing data, the PIAAC design followed a similar procedure to those used in prior studies (ALL and IALS) in order to provide comparability. Because this was a voluntary survey of the adult population without direct consequence to the test taker, missing data in PIAAC has a characteristic structure that relates to the matrix sampling design and the instituted accommodation for respondents with very low literacy skills through core items. This structure is in part characterized by data missing completely at random (MCAR; within each path due to random assignment of blocks) as well as data missing at random (MAR), due to the self-assigned choice of the paper versus computer path or the selection of this path based on background data. More specifically, there are different types of missing values within the *cognitive part* of PIAAC:

1) Missing by design: items that were not presented to each respondent due to the matrix sampling design used in PIAAC (see Chapter 1). Accordingly, these structural missing data, unrelated to respondents’ literacy, numeracy, and problem-solving skills, were ignored when calculating respondent proficiencies.

2) Omitted responses: missing responses that occurred when respondents chose not to perform one or more presented items, either because they were unable to do so or some other reason. Any missing response followed by a valid response (whether correct or incorrect) was defined as an omitted response. Omitted responses in the PBA were treated as wrong, because a random response to an open-ended item would almost certainly result in a wrong answer. In the case of the CBA, where it was possible to assess response times per item, nonresponses due to rapid omission were differentiated

from nonresponses after interaction with the stimuli (based on literature on response latencies; cf. Setzer & Allspach, 2007; Wise & DeMars, 2005; Wise & Kong, 2005). Thus, omitted responses were only treated as wrong if a responded spent more than five seconds on an item. If a respondent spent less than five seconds, the nonresponse was considered not attempted and treated as a missing value.

3) Not reached or not attempted responses: missing responses at the end of a block were treated as if they were not presented due to the difficulty of determining if the respondent was unable to finish these items or simply abandoned them.

Cases where respondents did not answer a sufficient number of background questions (< 5 items) were considered as incomplete cases and not used in the latent regression, and also not included in computing plausible values.

Some respondents who answered a sufficient number of background questions may not have been able to respond to the cognitive items or were unwilling to respond to the cognitive items. In these instances, the interviewers were required to document the extent to which the background questions and cognitive items were answered and to ascertain the reason for missing responses. These reasons may be categorized as:

1) Nonresponse due to refusal to participate, thus unrelated to literacy, numeracy, and problem-solving skills

2) Unable to respond due to a language difficulty or cognitive skill-related disability, thus indicating a deficiency of literacy, numeracy and problem-solving skills

3) Inability to provide a written response due to a physical disability

4) Other unspecified reasons

Only the missing responses of nonrespondents in the second category were imputed as incorrect. The rest of the missing responses were considered unrelated to cognitive skills and thus ignored.

On average across countries (based on the weighted and standardized data), 96.9% of respondents completed a BQ and responded to the cognitive items.

Respondents who correctly solved fewer than three of the six core items on the CBA, and fewer than four of the eight core items on the PBA (after the BQ and before the cognitive assessment) were not required to continue with an additional task booklet of cognitive items; their missing responses were considered incorrect for the proficiency estimation. This decision was based on the findings in the Field Test, which showed that respondents who correctly answered fewer than three of the six, or four of the eight core items, were not likely to provide a correct answer to more than 8% of items.

Treatment of respondents with fewer than five cognitive item responses

This section addresses the issue of respondents who provided background information but did not completely respond to the cognitive items. A minimum of five completed items per domain was necessary to assure sufficient information about the proficiency of respondents. On average, 1.7% of the PIAAC samples responded to fewer than five cognitive items per subscale.

Many large-scale assessment programs such as the National Assessment of Educational Progress (NAEP), the National Educational Longitudinal Study (NELS), and the 1985 Young Adult Literacy Survey (YALS) have excluded nonresponding cases from the analyses. Even though a proportion of the missing data and some of the characteristics of the missing data sample were reported, their impact on the analyses was not determined. This practice can yield both biased and inaccurate proficiency distributions for some subpopulations because of differential response rates among subpopulations. For example, individuals who were excluded based on a failure to answer core items for the 1985 YALS were predominantly Hispanic; hence, Hispanic subpopulation results were based only on those who read English, The summary table does not indicate the impact of the non-English readers within the Hispanic population. It should be emphasized again that the presence of extensive background information related to one's cognitive skill is necessary to implement any method for the imputation of proficiency scores.

In some cases, a sampled individual decided to stop the assessment. The reasons for stopping may be classified into two groups: those unable to respond to the cognitive items (i.e., for cognitive-related reasons), and those unwilling to respond (i.e., for noncognitive-related reasons). It should be noted that 2.8% of cognitive-related reasons were either "failed PBA core items" or "failed CBA core items."

PIAAC followed the ALL and IALS procedure with respect to cases with responses to fewer than five cognitive items per domain. All consecutively missing responses at the end of a block of items were treated as incorrect if the reason for not responding to the cognitive items was related to the cognitive skills (literacy, numeracy, problem solving). Otherwise, all consecutively missing responses were treated as "not reached."

This scoring method is important with regard to the latent regression population model described in section 17.2. The population model is used to estimate proficiency values based on responses to the background questions and the cognitive items. A respondent's proficiency is determined from an a posteriori distribution that is the product of two functions: a conditional distribution of proficiency given responses to the background questions, and a likelihood function of proficiency given responses to the cognitive items. The treatment of nonresponding examinees due to noncognitive-related reasons has no impact on the likelihood function of proficiency. On the other hand, there is an impact associated with the treatment for nonresponding cases due to cognitive-related reasons. In the latter case, the likelihood function will be very peaked at the lower end of the scale, which is believed to correctly represent the proficiency of those who are unable to respond to the cognitive items. With this scoring procedure, summary statistics can be produced for the entire population, including those who respond to cognitive items correctly in various degrees, as well as those who were not able to respond to cognitive items.

Furthermore, examinees with responses to fewer than five cognitive items per domain were not included in a first run of the population modeling (with regard to the regression model) to obtain unbiased Γ and Σ. In a second analysis, the regression parameters were treated as fixed to obtain plausible values for all cases, including those with fewer than five responses to cognitive items. More detailed information is provided in section 17.3.4.

Item statistics under adaptive testing

Nonadaptive large-scale population surveys such as Programme for International Student Assessment and Trends in International Mathematics and Science Study, where each block of items are administered to randomly equivalent respondents through a type of balanced incomplete block design, the standard item statistics represent entire samples. Solely based on this randomly equivalent groups responding to every item, the item statistics are comparable across items within a country as well as across countries. In comparison, PIAAC used two levels of adaptive testing resulting in that standard item statistics represent only subsets of the entire sample and these subsets were defined through type of skills and proficiencies. Thus the standard item statistics are not comparable across items within a country or across countries.

The first level of adaptation used in PIAAC is in terms of mode of administration. Through a series of questions and responses to the CBA core items, PBA items were administered to those without ICT skills and those who were not willing to participate in the CBA. The rest of the respondents in each country (those with ICT skills who were willing to take the assessment on the computer) took CBA. The proportions of the two groups differ by country and demographic characteristics such as age and education, and also they differ by ability. PBA and CBA items were not administered to randomly equivalent group of respondents.

The second level of adaptation in PIAAC was within the CBA portion of the assessment. PIAAC used a probability-based multistage adaptive algorithm where the cognitive items for literacy and numeracy were not administered to randomly equivalent groups of respondents. In other words, more able respondents received a more difficult set of items than less able respondents. Thus item statistics of “easy items” were no longer comparable with “difficult items.” Moreover, the countries differed in the distributions of skills, resulting in the distributions of administered items being different. CBA items were not administered to randomly equivalent group of respondents.

However, the comparability of item statistics across countries could be increased by standardizing the proportions of adaptive paths. Such an approach was used to evaluate block order effect in the next section.

Block order effect in the CBA

A block order effect is present when a different order of blocks of items impacts the proportion of correct item responses, that is, the item difficulty or some other characteristic of the item. Stated differently, examinee proficiency (with regard to the measured domains) and the manner in which the survey is administered influences the survey outcomes. As a precaution, the PIAAC design in the CBA was created in order to counterbalance the potential effects of item order on the difficulty of the items. In PIAAC, each respondent received two cognitive modules, where each module comprised either literacy, numeracy or problem-solving items. Each module of literacy and numeracy items appeared in two different positions within the assessment (block-order design: literacy – numeracy; numeracy – literacy, literacy – problem solving2; problem solving1 – literacy; numeracy – problem solving2; problem solving1 – numeracy; problem solving1 – problem solving2; see Chapter 1). The order of content-related blocks was examined to determine if there was any effect on the outcome of the literacy and numeracy proficiencies (note that it was not possible to examine order effects on the domain of problem solving in technology-rich environments as the different problem-solving blocks comprised different items, in contrast to the two other domains). Table 17.4 shows the average proportion correct for items in a given block for PIAAC; the average proportion is calculated from the weighted and standardized data for all participating countries. While the average proportions correct across all countries are virtually identical within 1 percentage point regardless of paired domains as long as domain order is the same, a slight block order effect was found, 2.8% for literacy modules and 1.3% for numeracy modules.

The weighted proportion correct for an item was calculated as follows:

$$P_i = \frac{\sum_k WP_k \sum_j W_j \left(x_{ji} = 1 \middle| k \right)}{\sum_k WP_k \left(\sum_j W_j \left(x_{ji} = 1 \middle| k \right) + \sum_j W_j \left(x_{ji} = 0 \middle| k \right) + \sum_j W_j \left(x_{ji} = 2 \middle| k \right) \right)}$$

where proportion correct on item i was calculated by using standardized weights of path k WP_k, final weights for the respondent j, scores responses correct ”1”, incorrect ”0”, and omit “2”.

Table 17.4: Average proportion correct; content-related block-by-block order (PIAAC Main Study)

Country	Average of Literacy Items 1st Module		Average of Numeracy Items 1st Module		Average of Literacy Items 2nd Module		Average of Numeracy Items 2nd Module	
	LIT-NUM	LIT-PS2	NUM-LIT	NUM-PS2	NUM-LIT	PS1-LIT	LIT-NUM	PS1-NUM
Australia	56.7%	58.8%	67.8%	67.2%	53.0%	55.9%	67.2%	67.1%
Austria	61.5%	61.0%	64.5%	65.1%	58.9%	58.8%	63.4%	63.1%
Canada	58.7%	58.4%	63.8%	62.5%	54.6%	55.6%	61.9%	62.4%
Cyprus[3]	49.4%		60.7%		45.8%		60.8%	
Czech Rep.	53.5%	54.4%	68.6%	65.4%	53.9%	51.6%	64.7%	66.5%
Denmark	58.7%	57.2%	68.9%	68.2%	55.0%	55.2%	67.0%	68.1%
England/N. Ireland (UK)	58.0%	57.6%	60.5%	60.8%	52.2%	51.8%	59.9%	60.4%
Estonia	57.0%	57.1%	65.7%	65.1%	54.2%	54.7%	65.4%	66.9%
Finland	65.5%	65.2%	72.5%	74.0%	63.3%	62.6%	70.2%	67.9%
Flanders (Belgium)	60.0%	57.9%	67.2%	69.7%	57.1%	58.5%	67.3%	65.5%
France	52.1%		60.2%		48.4%		58.8%	
Germany	57.1%	56.6%	66.3%	67.5%	53.0%	51.9%	65.9%	65.3%
Ireland	56.3%	56.4%	60.7%	60.9%	52.1%	50.7%	58.9%	56.5%
Italy	47.5%		56.9%		44.2%		55.6%	
Japan	67.0%	68.9%	75.7%	76.1%	64.3%	64.1%	73.9%	74.1%
Korea	57.2%	57.1%	62.9%	63.4%	56.9%	57.8%	62.9%	60.6%
Netherlands	62.8%	62.3%	68.5%	69.3%	59.6%	61.1%	69.0%	66.8%
Norway	60.3%	61.0%	69.2%	68.2%	59.1%	57.2%	66.2%	68.9%
Poland	56.6%	55.9%	61.5%	60.8%	51.3%	54.2%	62.1%	60.2%
Russian Fed.[4]	53.7%	52.9%	56.5%	58.4%	52.5%	50.4%	57.5%	56.0%
Slovak Rep.	54.5%	55.4%	67.2%	66.9%	53.8%	53.9%	67.0%	66.7%
Spain	48.4%		55.7%		44.8%		55.4%	
Sweden	62.4%	64.7%	69.7%	70.6%	58.5%	61.9%	67.0%	68.9%
United States	57.8%	56.7%	56.9%	58.8%	52.1%	54.9%	56.8%	55.0%
Average$_1$	58.8%	58.8%	65.7%	65.9%	55.8%	56.1%	64.7%	64.3%
Average$_2$	49.4%		58.4%		45.8%		57.7%	

Average$_1$ is based on the countries that participated in the problem solving domain.
Average$_2$ is based on the countries that did not participated in the problem solving domain.

17.3.2 National and international item calibration

Item calibration is the first step in population modeling and provides the item parameters for the cognitive items that are needed as one of the inputs for the population model used to calculate

[3] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[4] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

the plausible values (see section 17.2.). All cognitive items were calibrated using the 2PL or the GPCM model using *mdltm* (von Davier, 2005) for multidimensional discrete latent traits models. The software provides marginal maximum likelihood estimates (MML) obtained using customary expectation-maximization methods (EM), with optional acceleration. Both IRT models are described in detail in section 17.2.

Of the 166 items used for PIAAC, 18 literacy and 17 numeracy items were used as linking items between PBA and CBA (this means those items were identical between PBA and CBA); therefore, PIAAC contained 131 unique items. In other words, 166 items were described by 131 sets of item parameters. The 131 unique items were calibrated together with 132 unique items from IALS and ALL (263 unique items in total; see Table 17.5). The 100 reading component items were not used for the IRT calibration; for those items, descriptive statistics were provided such as percentage of correct responses, as well as overall timing of the reading component test (only 23.5% of the tested population received the reading component assessment). The 76 literacy items (described by 58 sets of item parameters), and the 76 numeracy items (described by 59 sets of item parameters) were scored dichotomously and calibrated using the 2PL in separate unidimensional IRT analyses. The 14 problem-solving items were scored dichotomously or polytomously and were calibrated using the 2PL and GPCM.

The item calibration also comprised a combined analysis using the IALS and ALL data for the purpose of producing linked scale for trend measurement (see section 17.4.2 and the IALS/ALL technical report for more details). Table 17.5 provides an overview of the distribution of the 263 unique cognitive items across the different surveys (ALL, IALS, PIAAC) and assessment modes (PBA, CBA).

Table 17.5: Distribution of the 263 unique cognitive items across surveys and assessment modes by domain used in PIAAC item calibration (Main Study)

		IALS only	**IALS + ALL**	**IALS + PIAAC**	**IALS + ALL + PIAAC**	**ALL only**	**ALL + PIAAC**	**PIAAC only**	**Total items in calib-ration**
Literacy	PBA	42	30	0	0	45	0	6	123
	CBA	0	0	1	5	0	6	22	34
	PBA+ CBA	0	0	0	3	0	15	0	18
Numeracy	PBA	0	0	0	0	12	0	10	22
	CBA	0	0	0	0	0	13	22	35
	PBA+ CBA	0	0	0	0	0	17	0	17
Problem solving	CBA	0	0	0	0	0	0	14	14
Total items in calibration		42	30	1	8	57	51	74	**263**

Note: Linking items are counted to avoid duplication.

Two out of the 24 countries participating in PIAAC (France and Russia[5]) were unable to meet the data delivery deadline due to organizational reasons. The data for these countries were not included in the item calibration to obtain the international item parameters. However, the data for these countries – after they were received – went through the same quality assurance and national item calibration (to provide national item parameters for items which showed deviation with regard to the international item parameters). Altogether, data from 154,714 PIAAC respondents were used for the international IRT calibration. During the item calibration, sample weights standardized to represent each country equally were used.

As the samples for each assessment (PIAAC, IALS, ALL) came from somewhat different populations with different characteristics, the calibration procedure needed to take into account the possibility of any systematic interaction between the samples and the items that were used to produce estimates of the item parameters and sample distributions. For this reason, a multiple-group IRT model was estimated using a mixture of normal population distributions (one for each sample) where item parameters were generally constrained to be equal across countries with a unique mean and variance for each country. The moments of these distributions were updated at each iteration during IRT calibration.

The item calibration was completed in two consecutive steps: First, the data were analyzed in an international calibration under the assumption that the common data (including the data from all participating countries) were comparable for all items in the assessment. This step was used to obtain estimates of the international (or common) item parameters, which were equal for all countries. In the subsequent step, national (or unique) item parameters were estimated in order to account for national deviations for a small subset of items. This involved a close monitoring of the IRT scaling for item-by-country interactions and allowing country-specific item parameters only in instances where substantial deviations were identified. An algorithmic approach that automatically identified those country-by-item combinations requiring national parameters based on DIF detection was applied. Items not exhibiting appropriate fit using an international parameter received a country-specific parameter. However, if more than one country exhibited a deviation from the international parameters, an algorithm was applied that ensured parsimony in the parameterization. For example, if two countries showed poor item fit for the same item in the international calibration, and in the same direction, both countries received the same unique item parameter estimated for these two countries (note that the term "national item parameters" in this report is used for both cases: one country that receives a unique country-specific item parameter, and more than one country that receive the same unique item parameter which is different from the international item parameter).

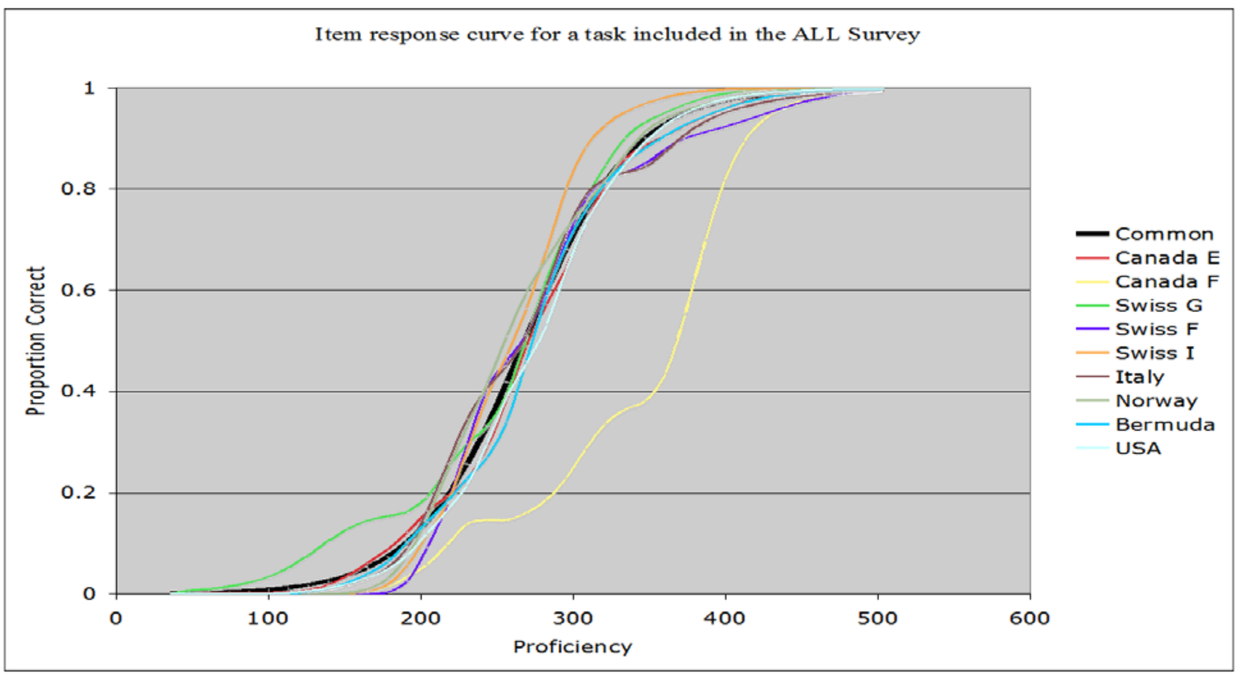
To identify misfitting items, fit statistics were estimated using the mean deviation (MD) and the root mean square deviation (RMSD). The MD is most sensitive to the difficulties of items and can represent a magnitude of shift of observed data from the estimated ICC. The RMSD is a standardized index of the discrepancy between the observed ICC and the model-based ICC; it is sensitive to measure the deviation of the observed item characteristics from the estimated ICC both in terms of slope and location of the item response function. Poorly fitting item characteristic curves were revealed using a RMSD > 0.1 criterion (a value of 0 indicates no discrepancy; in other words, a perfect fit of the model). The identification of poor fitting items and the replacement of international item parameters with country-specific (unique) parameters

[5] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

was carried out using an automatic algorithm in *mdltm*. Thus, the international and national calibrations were conducted simultaneously for all countries, that is, all estimated item parameters (international and national) are located on one common scale.

In most cases, the item responses across countries were accurately described by the international (common) item parameters. For some items, there was evidence that the estimated parameters did not fit as well for a certain assessment sample from a few countries as compared to the others. However, this pattern was not consistent for any one particular country. Given this estimation and optimization approach, no item was dropped from the analysis in PIAAC. For those items with item functions showing substantial deviation from the international item parameters (poor fitting items), national (unique) item parameters were estimated. If an item showed poor fit but had the same kind of poor fit in multiple countries, an additional country-group specific parameter besides the international or common item parameter was used for this item. If an item showed poor fit in one or two countries only or showed item fit to a different extent in different countries (unique deviation), the unique country-specific item parameters were used for further analysis. Thus, PIAAC allowed for different sets of item parameters to improve model fit and optimize the comparability of countries. Figure 17.1 shows a typical plot of a case (for the 2PL) to illustrate how the data from one country might not support the use of international item parameters.

Figure 17.1: Item response curve for an item where the international item parameter is not appropriate for one country



The solid black line is the fitted two-parameter logistic item response curve that corresponds to the international item parameters; the other lines are observed proportions of correct responses at various points along the proficiency scale for the data from each subpopulation. The horizontal axis represents the proficiency scale. This plot indicates that the observed proportions of correct responses, given the proficiency, are quite similar for most countries. However, the data for one country indicated by the yellow line shows a noticeable departure from the common ICC. This item is far more difficult in that particular country than expected given the responses on other items. Thus, a unique set of item parameters was estimated for that country.

Table 17.6 provides an overview of the number of country-specific (national) item parameters per country (see also Appendix 17.1 for detailed information), which were used together with the international parameters for the remainder of the items to calculate plausible values in PIAAC. For literacy, country-specific item parameters were estimated for only 8% of the items due to item-by-country interactions. For numeracy, 7% of the items necessitated country-specific parameters, and for problem solving, 3% of unique item parameters were used. (Unique item parameters for Russia[6] were determined after the reduction of the Russian sample by more than 1,200 cases due to issues in those data.)

Table 17.6: Number of national item parameters for each country and proficiency scale

Country	**Number of Country-Specific Item Parameters**	**Number of Country-Specific Item Parameters**	**Number of Country-Specific Item Parameters**
	Literacy (76 items)	**Numeracy (76 items)**	**Problem Solving (14 items)**
Australia	7	1	0
Austria	2	2	0
Canada (English)	2	1	0
Canada (French)	6	3	0
Cyprus[7]	11	3	NA
Czech Republic	8	5	1
Denmark	3	5	0
England/N. Ireland (UK)	3	3	0
Estonia	5	4	1
Finland	6	7	0
Flanders (Belgium)	5	5	0
France	8	3	NA
Germany	3	2	0
Ireland	2	2	0
Italy	5	3	NA
Japan	14	16	1
Korea	16	16	2
Netherlands	3	5	1
Norway	7	9	0
Poland	6	6	0
Russian Federation[8]	12	12	3
Slovak Republic	9	3	2
Spain	2	3	NA
Sweden	7	5	0
United States	5	9	0

[6] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

[7] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

[8] Please refer to the above note regarding the Russian Federation.

17.3.3 National reports

For the purposes of secondary analyses and transparency, every participating country received the prepared data files including plausible values for the international data, and the country-specific data, respectively. The reported values are based on the international calibration providing a common, comparable scale, with the potential adjustment of utilizing country specific item parameters to improve model fit and reduce bias. National reporting is supported by supplying these databases to each country, and additionally providing a set of tools for further analysis.

17.3.4 Generating plausible values

Plausible values are multiple imputed proficiency values based on information from the test items (the actual PIAAC literacy, numeracy, and problem solving tests) and information provided by the respondent in the BQ. Plausible values are used to obtain more accurate estimates of group proficiency than would be obtained through an aggregation of point estimates. A more detailed description is given in section 17.2 as well as in Mislevy (1991), Thomas (2002), and von Davier, Sinharay, Oranje & Beaton (2006).

In PIAAC, the computation of group-level reporting statistics involving scores in the three domains is based on 10 independently drawn plausible values for each scale assigned to each respondent. Each set of plausible values is equally well designed to estimate population parameters, however, multiple plausible values are required to represent the uncertainty in the domain measures appropriately (von Davier, Gonzalez & Mislevy, 2009). As mentioned earlier, the statistics based on scores are always computed at population or subpopulation levels. They should never be used to draw inferences at the individual level (see also section 18.4). Detailed information on the computation of plausible values in PIAAC is given in section 17.2.2.

For the population modeling and the calculation of plausible values for the scales of PIAAC, the computer program DGROUP (Rogers et al., 2006)[9] was used.

In the analyses of PIAAC, a normal multivariate distribution was assumed for $P(\theta_j|x_j, y_j, \Gamma, \Sigma)$, with a common variance, Σ, and with a mean given by a linear model with slope parameters, Γ, based on the principal components of several hundred selected main effects from the vector of background variables.

The item parameters for the cognitive items were obtained from the concurrent item calibration (see section 17.3.2) using the data from IALS, ALL and PIAAC as described above. The result of the concurrent calibration is a scale that provides comparable results across IALS, ALL and PIAAC. To calculate the plausible values for PIAAC only, the item parameters for the 166 PIAAC items (from the concurrent item calibration) were used in the population modeling.

The background variables included demographic information, educational experiences, occupational experiences and skill use, among others. A description of the different sections of the background data can be found in Chapter 3 of this report. All variables in the BQ were contrast coded before they were processed further in the population model. Contrast coding allows the inclusion of codes for refused responses as well as codes for responses that were not

[9] The statistical program DGROUP can be obtained from ETS on demand.

collected by means of routing and avoiding the necessity of linear coding. The increased number of variables obtained through contrast coding is substantial. To capture most of the common variance in the contrast-coded background questions with a reduced set of variables, a principal component analysis was conducted. Because each population can have unique associations among the background variables, a single set of principal components was not sufficient for all countries included in PIAAC. Therefore, the extraction of principal components was carried out separately by country. In PIAAC each set of principal components y^{c} (or conditioning variables) was selected to include 80 percent of the variance with the aim of explaining as much variance as possible while at the same time avoiding overparameterization.

Principal component scores based on nearly all background variables were used in PIAAC including international variables (collected by every participating country) as well as national background variables (country specific variables in addition to the international variables). Note, that the principal component analysis and the population modeling were calculated separately for each country in order to take into account the differences in associations between the background variables and the cognitive skills.

A small subset of respondents did not attempt the cognitive items or responded to fewer than five cognitive items for an inability to read or write in the language of assessment, a physical disability, a mental disability, or a refusal to participate in the survey. If these respondents had been excluded from the survey, the proficiency scores of some subpopulations in the PIAAC survey would have been systematically overestimated and the picture of the nation's cognitive skills would have been distorted. Those respondents with an insufficient number of responses (<5) to the cognitive items were excluded from the estimation of the latent regression. In a subsequent step, however, the latent linear regression estimated on the sample for examinees with sufficient numbers of responses was fixed and plausible values were drawn for all respondents. That is, in the second run all cases were included in the analysis but Γ and Σ were fixed to the values of the first run. Hence, a set of plausible values for the cognitive scales were calculated for all respondents regardless of the number of items attempted. The reason for this procedure is that sufficient information about the proficiency cannot be obtained for cases with fewer than five responses to cognitive items. Including these cases could influence the regression analysis, which aims to link background variables and (sufficiently accurate) proficiency estimates with the aim of predicting proficiency. For 2,616 cases across the 23 countries did not receive plausible values because of insufficient information due to literacy-related nonresponse.

17.4 Linking scales across delivery modes and surveys

PIAAC followed two aims with regard to the linking design:

1) Linking the different booklets containing different sets of items administered through different assessment (delivery) modes to each other in order to get comparable cognitive measures;
2) Linking the different ILSA adult surveys (IALS, ALL, PIAAC) to each other to provide trend measures.

17.4.1 Linking different booklets and assessment modes within PIAAC

To obtain comparable test results in all three cognitive domains for all sample groups, it was important that all items (in a given domain) were calibrated on one common scale. However, this was not easy to achieve given the complex test design in PIAAC. As illustrated in Chapter 1, PIAAC used a matrix sampling design where different items from the total item pool were administered to different test takers or groups by using different test booklets. Furthermore, items were administered through a version of adaptive testing, and by using different assessment modes, which made the design even more complex.

To establish a common scale for all items in a given domain, the items had to be linked together across test booklets (subset of items) and assessment modes. This was achieved by using common sets of items in the different booklets and assessment modes. Thus, certain items were administered in both the PBA and CBA (note that this pertains to literacy and numeracy items, as problem solving was only available for the CBA) as well as in different booklets (across different assessment modes). Out of 52 literacy and 52 numeracy items in the CBA, 18 literacy and 20 numeracy items were used to link the computer- and paper-based instruments. Within the CBA, all items were linked together in the booklet design. According to the distribution of the linking items, it was considered that the different item contexts (such as education, personal, work and everyday life), different item contents (such as data and chance, dimension and shape, quantity and number) and different cognitive processes or types of responses (such as integrate and interpret, evaluate and reflect, identify, and locate or access) were present within the linking items. In other words, the linking items were selected with the aim of being representative of the total item pool.

Through these linking items it was possible to calibrate items answered by different respondents in different booklets and assessment modes on one common scale for each cognitive domain. This was done within the item calibration (see section 17.3.2.). Deviations of item-by-country interactions were identified using a measure of MD and RMSD. Results for the PIAAC linking across assessment modes in the Main Study are presented in section 18.4.

17.4.2 Linking previous international adult assessments with PIAAC

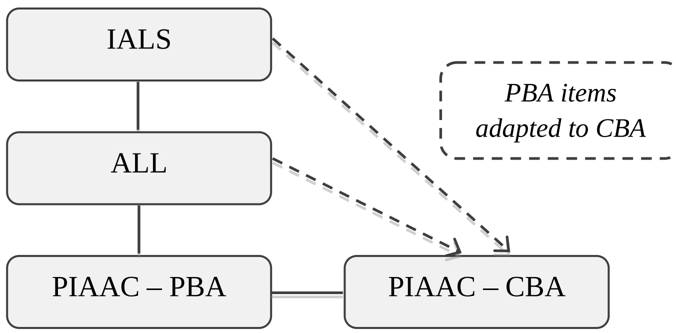
As the intent of PIAAC was to have its results linked to previous international adult assessments, 60 items of the literacy and numeracy items administered in PIAAC came from ALL and IALS. Seventy-four new items were developed for the literacy and numeracy domains, and new measures were developed for the reading components and problem solving domains (based on their respective frameworks) and tested in the PIAAC Field Test. Table 17.5 gives an overview of the item numbers per survey, domain and assessment mode.

The equivalence of item parameters among linking items from IALS and ALL to PIAAC was again evaluated through item calibration by applying IRT models (similar to the evaluation of the link between PBA and CBA in PIAAC).

Entire literacy items, including those unique to a particular survey as well as linking to multiple surveys, were reestimated using the entire aggregate data of IALS and ALL because the literacy scale in PIAAC is a joint scale of prose and document literacy scales (in IALS and ALL). These new parameters were used for the subsequent analyses. The numeracy scale was introduced in the ALL survey, and subsequent analyses used ALL numeracy item parameters.

Equivalence of item characteristics among the literacy and numeracy items common to IALS and ALL on the PBA was examined. As some IALS and ALL items (which used PBA only) were adapted to the CBA in PIAAC (see Figure 17.2), the equivalence of these adapted items to the appropriate IALS/ALL items was evaluated as well in the Field Test. Results for the PIAAC linking across surveys in the Main Study are presented in section 18.4.

Figure 17.2: Linking different international adult assessments and assessment modes (PIAAC)



To place the IALS and ALL items on the same scale as the PIAAC items, the item calibration (and thereby the linking) was used for the items and data from all three surveys. Therefore, the new estimates had to be transformed in order to be comparable to the old estimates, thus allowing the measurement of trend.

After the joint item calibration for all surveys was carried out, a linear transformation of the group means was conducted. The group means and standard deviations of the weighted scores obtained from the old item calibration of the IALS and ALL data were used to transform the new group means and standard deviations from the new joint item calibration (for IALS, ALL and PIAAC). An example of such a transformation is given in Table 17.7.

Table 17.7: Example of a transformation of IRT-based means of a set of old and new countries, calibrated together to find a transformation of the "new" countries' scores to the original scale

Old Countries	Original Mean	IRT New Calibration Based Mean	Transformed New Mean
A	240	0.3	240
B	250	0.4	250
C	260	0.5	260
D	270	0.6	270
E	280	0.7	280
New Countries	**Not Tested**		
F	-	0.3	240
G	-	0.5	260
H	-	0.7	280
I	-	0.55	265

For the trend measure, the transformed means of the weighted scores obtained from the item calibration were used for further analysis. The plausible values were influenced by this transformation as well but are not used for measuring trends.

References

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F.M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Johnson, E. G. (1989). Considerations and techniques for the analysis of NAEP data. *Journal of Educational Statistics, 14(4)*, 303-334.

Johnson, E. G., & Rust, K. F. (1992). Population inferences and variance estimation for NAEP data. *Journal of Educational Statistics, 17*, 175-190.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Mislevy, R. J. (1985). Estimation of latent group effects. J*ournal of the American Statistical Association, 80*(392), 993–997.

Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika, 56*(2), 177–196.

Mislevy, R. J., & Sheehan, K. M. (1987). Marginal estimation procedures. In A. E. Beaton (Ed.), Implementing the new design: The NAEP 1983-84 technical report. (Report No. 15-TR-20). Princeton, NJ: Educational Testing Service.

Muraki, E. (1992). A generalized partial credit model: application of an EM algorithm. *Applied Psychological Measurement, 16*(2), 159–177.

Rogers, A., Tang, C., Lin, M.-J., & Kandathil, M. (2006). DGROUP (computer software). Princeton, NJ: Educational Testing Service.

Setzer, J. C., & Allspach, J. R. (2007, October). *Studying the effect of rapid guessing on a low-stakes test: An application of the effort-moderated IRT model.* Paper presented at the annual meeting of the Northeastern Educational Research Association, Rocky Hill, CT. http://www.psyc.jmu.edu/assessment/research/pdfs/SetzerAllspach_NERA07.pdf

Sukkarieh, J., von Davier, M. & Yamamoto, K. (2012). *From biology to education: Scoring and clustering multilingual text sequences and other sequential tasks* (Research Report Series. ETS RR–12-25

Thomas, N. (1993). Asymptotic corrections for multivariate posterior moments with factored likelihood functions. *Journal of Computational and Graphical Statistics, 2*, 309–322.

Thomas, N. (2002). The role of secondary covariates when estimating latent trait population distributions. *Psychometrika, 67*(1), 33-48.

von Davier, M. (2005). *A general diagnostic model applied to language testing data.* Research Report RR-05-16). Princeton, NJ: Educational Testing Service.

von Davier, M., Gonzalez, E. & Mislevy, R. (2009) What are plausible values and why are they useful? *In: IERI Monograph Series: Issues and Methodologies in Large Scale*

Assessments, Vol. 2. Retrieved from IERI website: http://www.ierinstitute.org/IERI_Monograph_Volume_02_Chapter_01.pdf

von Davier, M. Sinharay, S., Oranje, A. & Beaton, A. (2006) Statistical procedures used in the National Assessment of Educational Progress (NAEP): Recent developments and future directions. In C. R. Rao and S. Sinharay (Eds.), *Handbook of Statistics (Vol. 26): Psychometrics*. Amsterdam: Elsevier.

Wingersky, M., Kaplan, B., & Beaton, A. E. (1987). Joint estimation procedures. In A. E. Beaton (Ed.), Implementing the new design: The NAEP 1983-84 technical report (pp. 285-292). Princeton, NJ: ETS.

Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment, 10*(1), 1-17.

Wise, S. L., & Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education, 18*, 163-183.

Appendixes

Appendix 17.1: Items per country that received country-specific item parameters in the population modeling

Item	Australia	Austria	Flanders (Belgium)	Canada (Eng.)	Canada (Fr.)	Cyprus	Czech Rep.	Germany	Denmark	France	Spain	Estonia	Finland	England/N. Ireland (UK)	Ireland	Italy	Japan	Korea	Netherlands	Norway	Poland	Russia	Slovak Rep.	Sweden	United States
LITERACY																									
C301C05S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C300C02S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
D302C02S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
D311701S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E321001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E321002S	*	*	*	*	*	**X**	*	*	*	*	*	*	*	*	*	*	*	*	*	Δ	**O**	*	*	*	*
C308117S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C308119S	*	*	*	**X**	*	*	Δ	*	*	*	*	*	*	*	**O**	**U**	*	*	*	*	*	*	*	*	*
C308120S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C308121S	*	*	*	*	*	*	*	*	**X**	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C305215S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C305218S	**X**	*	*	**O**	**X**	*	*	*	*	**V**	*	*	*	*	*	*	Δ	*	**U**	Δ	*	*	*	*	*
D315512S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	**X**	*	*	*	Δ	*	*	*	*
C308118S	*	**O**	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	**X**	**X**	**U**	Δ	*	*
D304710S	*	*	*	*	*	**X**	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
D304711S	*	*	*	*	**X**	Δ	*	*	*	Δ	*	*	*	*	*	*	**X**	*	*	*	Δ	Δ	*	*	*
C308116S	*	**X**	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E327001S	*	*	*	*	**X**	*	*	*	*	*	*	*	*	*	*	Δ	**O**	*	*	*	*	**U**	*	*	*
E327002S	*	*	*	*	*	**X**	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Δ	*	*	*	*
E327003S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Δ	*	**X**	*	**X**	*	*

Item	Australia	Austria	Flanders (Belgium)	Canada (Eng.)	Canada (Fr.)	Cyprus	Czech Rep.	Germany	Denmark	France	Spain	Estonia	Finland	England/N. Ireland (UK)	Ireland	Italy	Japan	Korea	Netherlands	Norway	Poland	Russia	Slovak Rep.	Sweden	United States
E327004S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*
D307401S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
D307402S	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C309319S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*
C309320S	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C309321S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C309322S	*	X	*	*	*	*	*	*	*	U	*	Δ	*	*	*	*	O	*	*	*	*	*	*	*	*
E322001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E322002S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*
E322005S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*
C313412S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C313414S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E322003S	*	*	*	Δ	X	*	*	*	*	O	*	*	X	X	*	*	*	*	*	*	*	*	*	*	X
C310406S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C310407S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E320001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*
E320003S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*
E320004S	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E322004S	*	*	*	X	*	*	*	*	*	O	*	*	*	*	*	Δ	*	*	*	*	*	*	*	*	*
D306110S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
D306111S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	X	*	*	*	*	*	*	*	*
C313410S	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	Δ	*	*	*	*	*	*	*
C313411S	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C313413S	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E323003S	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Item	Australia	Austria	Flanders (Belgium)	Canada (Eng.)	Canada (Fr.)	Cyprus	Czech Rep.	Germany	Denmark	France	Spain	Estonia	Finland	England/N. Ireland (UK)	Ireland	Italy	Japan	Korea	Netherlands	Norway	Poland	Russia	Slovak Rep.	Sweden	United States
E323004S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	Δ	*	*	*
E318001S	*	*	*	*	*	*	*	X	*	*	Δ	*	*	*	*	X	*	*	X	O	*	U	*	*	*
E318003S	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E329002S	*	*	O	*	*	*	*	*	*	*	*	X	Δ	Δ	*	X	*	*	*	U	*	*	*	O	Δ
E329003S	*	*	*	*	*	*	*	O	*	*	X	*	*	*	*	*	Δ	*	Δ	X	*	V	U	*	*
E323002S	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	X	Δ	*	*	*	*	*	*	*	*
E323005S	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*
M301C05S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
P330001S	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Δ	*	*	*	*	*	O	*	*
N302C02S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
M300C02S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*
N306110S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*
N306111S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*
M313410S	X	Δ	*	*	*	*	*	*	*	O	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
M313411S	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	Δ	*	*	*
M313412S	X	*	*	*	*	*	Δ	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*
M313413S	*	X	*	*	*	*	*	*	*	*	*	Δ	*	*	*	O	X	*	*	*	U	*	*	X	*
M313414S	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	Δ	*	*	*	*	*	*	*	*	*
P324002S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*
P324003S	X	*	*	*	Δ	*	*	*	*	*	*	O	*	*	*	U	*	*	*	*	*	*	*	*	*
M305215S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
M305218S	*	*	*	*	*	*	*	*	*	U	Δ	X	*	*	*	*	*	*	X	*	Δ	O	X	*	*

Item	Australia	Austria	Flanders (Belgium)	Canada (Eng.)	Canada (Fr.)	Cyprus	Czech Rep.	Germany	Denmark	France	Spain	Estonia	Finland	England/N. Ireland (UK)	Ireland	Italy	Japan	Korea	Netherlands	Norway	Poland	Russia	Slovak Rep.	Sweden	United States
P317001S	Δ	*	X	X	*	*	*	*	*	*	*	Δ	*	*	*	*	O	*	*	*	*	*	*	*	*
P317002S	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	Δ	*	*	*	X	O	*	*	*
P317003S	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
M310406S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*
M310407S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
M309319S	Δ	*	*	*	*	*	Δ	*	*	O	*	*	*	*	X	*	*	X	X	*	*	O	*	*	*
M309320S	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*
M309321S	*	*	*	*	X	*	*	*	*	Δ	*	*	*	*	X	*	*	*	*	*	*	Δ	*	O	*
M309322S	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	X	X	*	*	*	*	*	*
NUMERACY																									
C600C04S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C601C06S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E645001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C615602S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*
C615603S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*
C624619S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*
C624620S	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	Δ	*	*	*	X	*	O	*	*	*
C604505S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C605506S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*
C605507S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	Δ	*	*	*
C605508S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*
E650001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	Δ	*	*	*

Item	Australia	Austria	Flanders (Belgium)	Canada (Eng.)	Canada (Fr.)	Cyprus	Czech Rep.	Germany	Denmark	France	Spain	Estonia	Finland	England/N. Ireland (UK)	Ireland	Italy	Japan	Korea	Netherlands	Norway	Poland	Russia	Slovak Rep.	Sweden	United States
C623616S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C623617S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*
E657001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Δ	X	*	*	X	*	O	*	*	*
C619609S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*
E632001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E632002S	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	Δ	X	*	*	*	*
E646002S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*
C620610S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C620612S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*
C613520S	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C614601S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C618607S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C618608S	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	Δ	*	*	*	*	*	O	*	*	*
E635001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C607510S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*
E655001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*
C602502S	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	Δ	*	*	*	*	O	*	*	*
C602503S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	Δ	*	*	*	*	O	*	*	*
C608513S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C602501S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C606509S	*	*	*	*	*	*	*	*	X	U	Δ	*	*	*	*	O	*	*	*	*	*	*	*	*	*
C611516S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	Δ	*	*	*
C611517S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C622615S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*
E665001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*

Item	Australia	Austria	Flanders (Belgium)	Canada (Eng.)	Canada (Fr.)	Cyprus	Czech Rep.	Germany	Denmark	France	Spain	Estonia	Finland	England/N. Ireland (UK)	Ireland	Italy	Japan	Korea	Netherlands	Norway	Poland	Russia	Slovak Rep.	Sweden	United States
E665002S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E636001S	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*	Δ	*	*	*
C617605S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	Δ
C617606S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*
E660003S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*
E660004S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E641001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	Δ	*	*	*
E661001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E661002S	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Δ	*	*	*	*	*
C612518S	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*	Δ	X	X	*
E651002S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	Δ	*	*	*	*	*	*	*	*
E664001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	Δ	*	O	*	*	*
E634001S	*	*	*	*	*	*	X	*	*	*	X	*	*	*	*	*	*	X	Δ	*	*	*	*	O	*
E634002S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*
E644002S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	Δ	*	*	*	*	*	*	*	*
M600C04S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	X	*
P601C06S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*	*
P614601S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*
P645001S	*	*	X	X	*	*	*	*	*	*	*	O	*	*	*	*	Δ	*	*	*	*	*	*	Δ	*
M615602S	*	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
M615603S	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	Δ	O	*	*	*	*	*	*	*	*
P640001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	X	*	*	*	*	*	*	*
M620610S	X	Δ	*	*	X	*	*	*	Δ	*	*	*	*	Δ	X	*	*	O	U	*	O	*	*	*	*
M620612S	*	*	*	*	*	*	*	*	*	Δ	*	*	*	*	X	*	*	*	*	*	*	*	*	*	*
P666001S	*	*	*	O	X	*	U	*	*	*	*	*	X	V	*	W	*	Δ	*	Δ	*	Z	*	*	Δ

Item	Australia	Austria	Flanders (Belgium)	Canada (Eng.)	Canada (Fr.)	Cyprus	Czech Rep.	Germany	Denmark	France	Spain	Estonia	Finland	England/N. Ireland (UK)	Ireland	Italy	Japan	Korea	Netherlands	Norway	Poland	Russia	Slovak Rep.	Sweden	United States
M623616S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
M623617S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	**X**	*	*	*	*	*	Δ	*	*	*
M623618S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	**X**	*	*	*
M624619S	*	**X**	*	*	*	*	*	*	*	*	*	Δ	*	*	*	*	*	**O**	*	*	*	*	*	*	*
M624620S	*	*	*	*	*	*	*	**X**	*	*	*	Δ	**X**	*	*	*	Δ	*	**X**	*	*	*	*	*	*
M618607S	*	*	*	**X**	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	**O**	Δ	*	*
M618608S	*	**O**	*	*	*	*	*	Δ	**X**	**U**	**X**	Δ	**X**	*	*	*	*	*	**X**	*	*	*	Δ	*	*
M604505S	*	*	*	*	*	**X**	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
M610515S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	**X**	*	*	*	*	*	*
P664001S	*	**X**	*	*	*	*	*	**X**	*	*	*	*	*	*	*	*	*	*	**X**	*	*	*	**X**	*	*
M602501S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
M602502S	*	*	*	*	*	**X**	*	*	*	*	*	*	*	*	*	*	*	*	Δ	*	*	*	**O**	*	*
M602503S	*	*	*	*	*	*	*	**X**	*	*	*	*	*	*	*	Δ	**O**	*	**U**	*	*	*	*	*	*
P655001S	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	**X**	*
PSTRE																									
U01A000P	*	*	*	*	#	*	*	*	#	#	*	*	*	*	#	*	*	*	*	*	*	*	*	*	*
U01B000S	*	*	*	*	#	*	*	*	#	#	*	*	*	*	#	*	*	*	*	*	*	*	*	*	*
U03A000S	*	*	*	*	#	*	*	*	#	#	*	*	*	*	#	*	*	*	*	*	Δ	**X**	*	*	*
U06A000S	*	*	*	*	#	*	*	*	#	#	**X**	*	*	*	#	**X**	*	*	*	*	*	Δ	*	*	*
U06B000S	*	*	*	*	#	**X**	*	*	#	#	*	*	*	*	#	*	*	*	*	*	Δ	*	*	*	*
U21X000S	*	*	*	*	#	*	*	*	#	#	*	*	*	*	#	*	*	*	*	*	*	*	*	*	*
U04A000P	*	*	*	*	#	*	*	*	#	#	*	*	*	*	#	*	*	*	*	*	*	*	*	*	*
U19A000S	*	*	*	*	#	*	*	*	#	#	*	*	*	*	#	*	**X**	*	*	*	*	*	*	*	*
U19B000P	*	*	*	*	#	*	*	*	#	#	*	*	*	*	#	*	*	**X**	*	*	*	*	*	*	*
U07X000S	*	*	*	*	#	*	*	*	#	#	*	*	*	*	#	*	*	*	*	*	*	**X**	*	*	*

Item	Australia	Austria	Flanders (Belgium)	Canada (Eng.)	Canada (Fr.)	Cyprus	Czech Rep.	Germany	Denmark	France	Spain	Estonia	Finland	England/N. Ireland (UK)	Ireland	Italy	Japan	Korea	Netherlands	Norway	Poland	Russia	Slovak Rep.	Sweden	United States
U02X000P	*	*	*	*	#	*	*	*	#	#	*	*	*	*	#	*	*	*	*	*	*	*	*	*	*
U16X000S	*	*	*	*	#	*	*	*	#	#	*	*	*	*	#	*	*	*	*	*	*	*	*	*	*
U11B000P	*	*	*	*	#	*	*	*	#	#	*	*	*	*	#	*	**X**	*	*	*	*	*	*	*	*
U23X000P	*	*	*	*	#	*	*	*	#	#	*	*	*	*	#	*	*	*	*	*	*	*	*	*	*

Note: * denotes international item parameters; all other symbols and letters (**X**, Δ, **O**, **U**, **V**, **W**, **Z**) denote country-specific item parameters; identical symbols/letters in the same row (or for the same item) for different countries denote identical item parameters for the specific item in these countries (identical symbols/letters in different rows/items do not); # denotes items that were not presented in a country – typically this symbol will be found for countries that optioned out of the assessment of PSTRE.

Chapter 18: Scaling Outcomes

Kentaro Yamamoto, Lale Khorramdel and Matthias von Davier

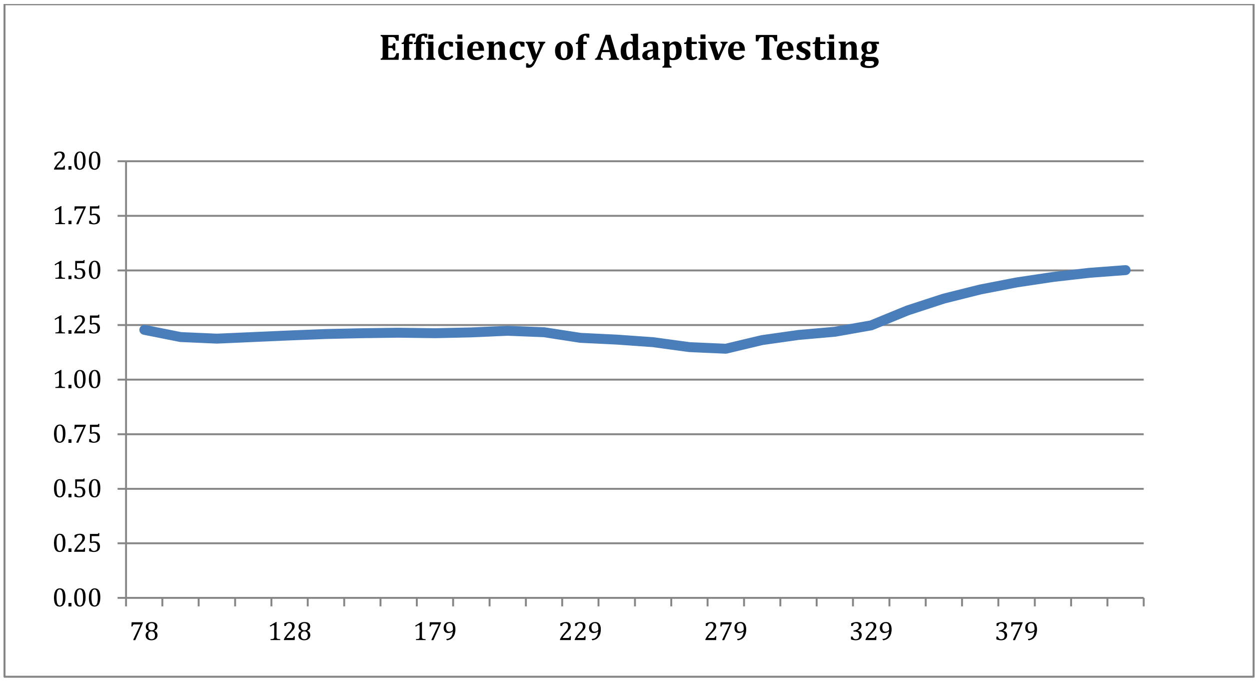
18.1 International characteristics of the PIAAC item pool and scales

18.1.1 Test information and evaluation of adaptive testing

The PIAAC multistage adaptive testing design for the CBA was developed to match a respondent's background profile and ability while maintaining a degree of randomness of assignment to ensure broad coverage of the domain in all proficiency levels. This made it possible to match respondents' abilities with the booklets' difficulties in a fair manner. Moreover, it was possible to control the exposure rates for all booklets (cf. Chen, Yamamoto & von Davier, in press). The aim of adaptive testing is to increase efficiency, validity and accuracy of the cognitive measurement. The multistage adaptive testing design may also increase engagement and test motivation, and hence reduce nonresponse and random responding.

The graph in Figure 18.1 shows the efficiency of the PIAAC multistage adaptive assessment for the literacy scale over averaged (expected) test information of the nonadaptive assessment, defined as the ratio of the conditional maximum test information of the 12 adaptive tests (note that one test consists of two clusters of items: stage 1 and stage 2) over the average test information of nonadaptive tests. The ratio of the two test information curves is shown on the vertical axis whereas the literacy scale is shown on the horizontal axis. Between the literacy proficiency values 100 and 400, the adaptive assessment was 15% to 47% more efficient than the average nonadaptive assessment based on the identical item set. Increased efficiency of adaptive testing means that the same amount of test information was obtained from the adaptive test as would be a nonadaptive test with 15-47% more items (or restated, the adaptive test required 13-32% fewer items).

Figure 18.1: Efficiency of the PIAAC multistage adaptive assessment for the scale of literacy over averaged (expected) test information of the nonadaptive assessment



18.1.2 Testing time

Each block of items for the domains of literacy, numeracy and PSTRE in the CBA was expected to take 30 minutes on average, including orientations. However, it turned out that in most cases, respondents took less than the expected amount of time (see Table 18.1). Table 18.4 shows the average time per item and block for the cognitive domains in PIAAC; the information in the tables does not include the average time spent for orientations. The reading components, which were expected to take 10 minutes on average, took less time as well (see Table 18.2).

Table 18.1: Average minutes per block of items in the CBA with regard to domain

Literacy		Numeracy		PSTRE	
Block	**Min (average)**	**Block**	**Min (average)**	**Block**	**Min (average)**
Core	1.19	Core	1.76	PS1 Block	20.65
CBA Block	22.52	CBA Block	21.89	PS2 Block	18.32

Table 18.2: Average time (in minutes) per block of items for the reading components domain

Block	Minutes (average)	SD
Vocabulary	2.48	1.86
Sentence	2.89	1.82
Passage 1, 2, 3, 4	6.30	3.64

* Vocabulary: 34 items. Sentence: 22 items. Passages 1, 2, 3 and 4: 44 items)

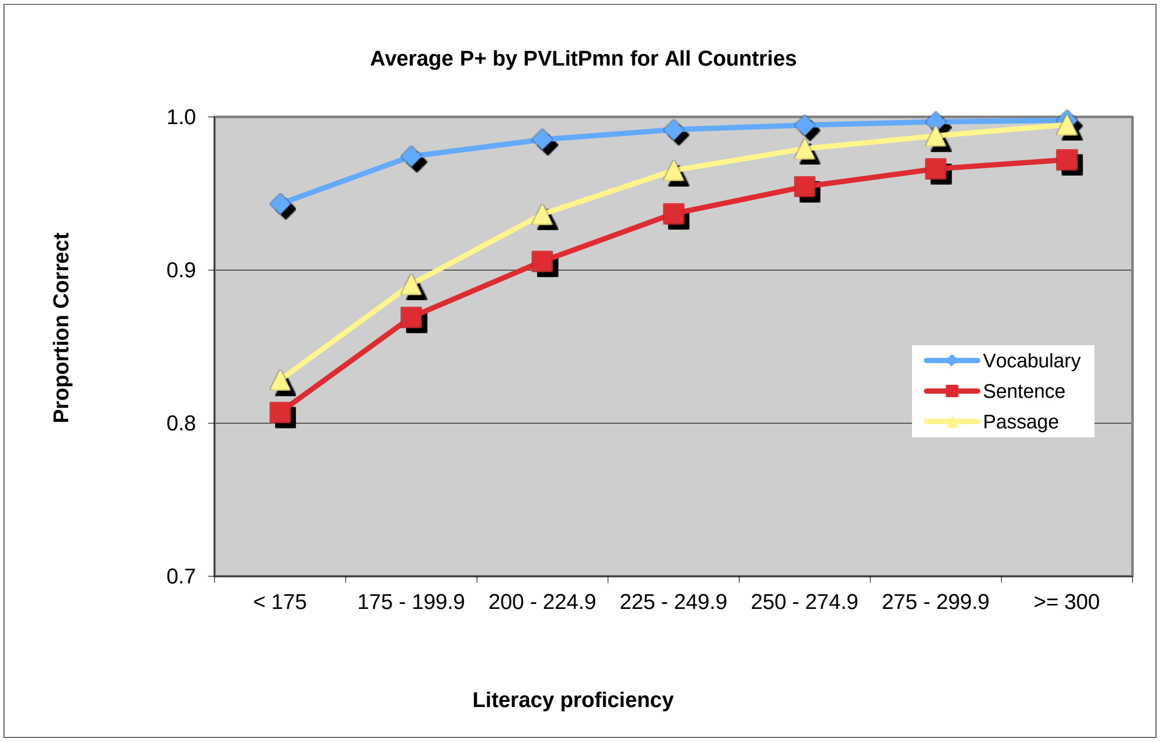
For the reading components, both response time and proportions correct had predictable relationships with literacy proficiencies (see Table 18.3). Results show a high proportion of correct responses as expected even among least able respondents of less than 175 (vocabulary: P+ = .94; sentence processing: P+ = .81; basic passage comprehension: P+ = .83), meaning the reading components were easy for every respondent. While high response accuracy was even among least able respondents, response fluency represented by the average response time indicate that less able respondents took 2.5 times longer to answer reading component items.

Table 18.3 Reading components average proportions correct and average response time by literacy posterior means

		Literacy Posterior Means						
		< 175	**175 - 199.9**	**200 - 224.9**	**225 - 249.9**	**250 - 274.9**	**275 - 299.9**	**>= 300**
Vocabulary	Average proportions correct P+	0.94	0.97	0.99	0.99	0.99	1.00	1.00
	Average response time per item (sec)	7.62	5.97	4.95	4.40	3.95	3.69	3.33
Sentence	Average proportions correct P+	0.81	0.87	0.91	0.94	0.95	0.97	0.97
	Average response time per item (sec)	14.45	10.99	9.23	8.19	7.24	6.69	6.06
Passage	Average proportions correct P+	0.83	0.89	0.94	0.97	0.98	0.99	0.99
	Average response time per item (sec)	16.63	13.29	10.93	9.26	8.18	7.33	6.38

Figure 18.2 shows the proportion of correct responses for the reading components scale projected onto the literacy proficiency scale.

Figure 18.2: Accuracy (discrimination by means of conditional P+) of the PIAAC scale for reading components projected onto the literacy proficiency scale, Main Study



18.1.3 Test reliability and accuracy

As different sets of items were administered to different respondents in the Main Study, it is not reasonable to calculate marginal reliabilities for each cognitive domain. In order to get an indication of test reliability, the explained variance for each cognitive domain (see Table 18.4) was computed based on the weighted posteriori variance. The explained variance shows how much variance is explained by the model; it is computed using the 10 plausible values as follows: 1 – (*expected error variance/total variance*). The weighted posteriori variance is an expression of the posterior measurement error and is obtained through the population modeling. The expected error variance is the weighted average of the posteriori variance. This term was estimated using the weighted average of the variance of the plausible values (the posteriori variance is the variance across the 10 plausible values). The total variance was estimated using a resampling approach (Efron, 1982). It was estimated for each country depending on the country-specific proficiency distributions for each cognitive domain.

Table 18.4: Test reliability for literacy, numeracy, and PSTRE

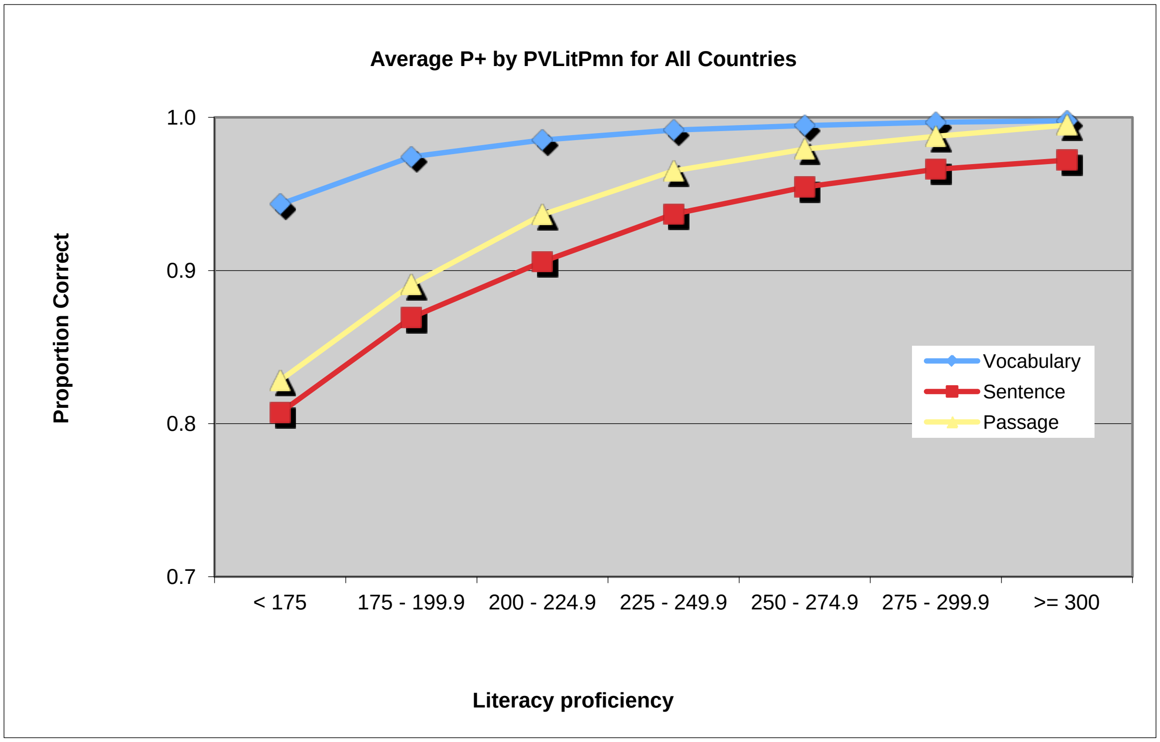
Countries	Literacy	Numeracy	PSTRE
Australia	0.879	0.875	0.834
Austria	0.865	0.860	0.844
Canada	0.878	0.874	0.847
Cyprus[1]	0.846	0.860	---
Czech Republic	0.856	0.862	0.869
Denmark	0.886	0.874	0.860
England/N. Ireland (UK)	0.879	0.896	0.876
Estonia	0.844	0.844	0.852
Finland	0.882	0.866	0.854
Flanders (Belgium)	0.881	0.868	0.846
France	0.890	0.902	---
Germany	0.887	0.889	0.864
Ireland	0.875	0.874	0.844
Italy	0.858	0.871	---
Japan	0.841	0.839	0.824
Korea	0.854	0.856	0.828
Netherlands	0.889	0.888	0.849
Norway	0.887	0.892	0.871
Poland	0.854	0.852	0.845
Russian Federation[2]	0.844	0.839	0.887
Slovak Republic	0.839	0.858	0.800
Spain	0.891	0.895	---
Sweden	0.903	0.903	0.886
United States	0.898	0.907	0.866

The table above shows that the explained variance by the combined IRT and latent regression model is at a comparable level across countries. While the joint model including background and item response data reaches levels of around 0.85, it is important to keep in mind that this is not to be confused with a classical reliability coefficient, as it is based on more than the item responses. Comparisons among individual respondents are not appropriate, because the apparent accuracy of the measures is obtained by statistically adjusting the estimates based on background data. This approach does provide improved behavior of subgroup estimates, while the plausible values obtained using this methodology are not suitable for comparisons of individuals (e.g., Mislevy & Sheehan, 1987; von Davier, Sinharay, Oranje, & Beaton, 2006).

The accuracy of the reading components in PIAAC was good as well. Results show a high proportion of correct responses as expected (vocabulary: P+ = 97.4; sentence processing: P+ = 91.4; basic passage comprehension: P+ = 92.7), meaning the reading components were easy for every respondent. Figure 18.3 shows the proportion of correct responses for the reading components scale projected onto the literacy proficiency scale.

[1] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[2] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Figure 18.3: Accuracy (discrimination by means of conditional P+) of the PIAAC scale for reading components projected onto the literacy proficiency scale, Main Study



18.1.4 Domain intercorrelations

The estimated correlations (corrected for attenuation) between the three PIAAC domains per country range from .737 to .875 (see Table 18.5). The correlations are rather high, as expected, but still show there is some distinction between each of the domains.

Table 18.5: Estimated average intercorrelations of the domains of literacy, numeracy and PSTRE by country, based on plausible values

Countries	Literacy with Numeracy	Literacy with PSTRE	Numeracy with PSTRE
Australia	0.890	0.801	0.729
Austria	0.863	0.791	0.714
Canada	0.868	0.813	0.740
Cyprus[3]	0.813	---	---
Czech Republic	0.798	0.768	0.697
Denmark	0.876	0.816	0.762
England/N. Ireland (UK)	0.875	0.773	0.769
Estonia	0.833	0.801	0.750
Finland	0.864	0.809	0.714
Flanders (Belgium)	0.873	0.811	0.734
France	0.863	---	---
Germany	0.872	0.806	0.753
Grand Total	0.861	0.781	0.725
Ireland	0.871	0.770	0.703
Italy	0.827	---	---
Japan	0.855	0.717	0.668
Korea	0.882	0.766	0.696
Netherlands	0.886	0.824	0.767
Norway	0.895	0.801	0.763
Poland	0.852	0.749	0.682
Russian Federation[4]	0.790	0.685	0.694
Slovak Republic	0.854	0.716	0.662
Spain	0.887	---	---
Sweden	0.893	0.791	0.746
United States	0.888	0.813	0.759

[3] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[4] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

18.2 Scaling outcomes

18.2.1 Conditioning (population modeling)

As described in sections 17.2, and 17.3, the conditioning (population modeling) combining IRT models and latent regression models utilized all the background data for each of the PIAAC countries. Analyses were carried out by country to allow country-specific latent regression models. The resulting estimates were then used to generate plausible values.

To assure the conditioning worked well across countries, examinations of convergence efficiency, residual variances and correlations based on the 10 plausible values were conducted for each country and cognitive scale (correlations were computed with each plausible value, then the average calculated). Results showed comparable correlations among scales (see Table 18.5), comparable levels of reliability (see Table 18.4), and reasonable correlations with skill use self-reports (see Table 18.6 for selected correlations and Appendix 18.1 for detailed information).

Table 18.6: Marginal correlations per country of the respective domains with selected scales of the BQ, based on the 10 plausible values obtained from the population modeling (conditioning)*

Countries	LIT – Use of reading skills at home	LIT – Use of reading skills at work	LIT – Use of writing skills at home	LIT – Use of writing skills at work	NUM – Use of NUM skills at home	NUM – Use of NUM skills at work	PSTRE – Use of ICT skills at home	PSTRE – Use of ICT skills at work
Australia	0.335	0.211	0.280	0.210	0.319	0.200	0.307	0.218
Austria	0.327	0.290	0.257	0.208	0.276	0.260	0.350	0.255
Canada	0.337	0.214	0.229	0.175	0.273	0.194	0.318	0.19
Cyprus[5]	0.168	0.099	0.094	0.121	0.144	0.146	----	----
Czech Rep.	0.324	0.196	0.195	0.172	0.230	0.200	0.295	0.178
Denmark	0.328	0.220	0.255	0.152	0.251	0.235	0.323	0.224
England/N. Ireland (UK)	0.314	0.271	0.253	0.209	0.269	0.205	0.368	0.289
Estonia	0.315	0.197	0.229	0.155	0.270	0.219	0.403	0.240
Finland	0.313	0.207	0.279	0.171	0.315	0.253	0.384	0.216
Flanders (Belgium)	0.343	0.322	0.238	0.218	0.265	0.269	0.376	0.288
France	0.395	0.344	0.272	0.230	0.338	0.279	----	----
Germany	0.368	0.262	0.230	0.158	0.329	0.252	0.373	0.235
Ireland	0.320	0.243	0.245	0.192	0.243	0.215	0.355	0.273
Italy	0.362	0.281	0.167	0.210	0.239	0.258	----	----
Japan	0.271	0.153	0.052	0.095	0.179	0.256	0.267	0.246
Korea	0.370	0.252	0.153	0.163	0.298	0.211	0.309	0.208
Netherlands	0.336	0.252	0.280	0.188	0.278	0.229	0.365	0.208
Norway	0.283	0.236	0.176	0.178	0.236	0.227	0.335	0.257
Poland	0.384	0.226	0.264	0.13	0.314	0.223	0.317	0.149

[5] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 18.6 (cont.): Marginal correlations per country of the respective domains with selected scales of the BQ, based on the 10 plausible values obtained from the population modeling (conditioning)*

Countries	LIT – Use of reading skills at home	LIT – Use of reading skills at work	LIT – Use of writing skills at home	LIT – Use of writing skills at work	NUM – Use of NUM skills at home	NUM – Use of NUM skills at work	PSTRE – Use of ICT skills at home	PSTRE – Use of ICT skills at work
Russian Fed.[6]	0.273	0.098	0.084	0.088	0.294	0.119	0.263	0.122
Slovak Rep.	0.370	0.193	0.135	0.139	0.278	0.173	0.188	0.150
Spain	0.393	0.273	0.261	0.224	0.291	0.215	----	----
Sweden	0.279	0.184	0.203	0.156	0.223	0.243	0.377	0.269
United States	0.245	0.184	0.200	0.137	0.235	0.141	0.333	0.220

Note: The correlations for the ICT scales might be underestimated as not every respondent received the ICT items according to the path of the adaptive testing.
*LIT = Literacy, NUM = Numeracy, PSTRE = Problem-solving in technology-rich environments

The conditioning model estimations converged without any apparent issues, the between-scale correlations across countries are similar, and the correlations of direct assessed proficiency data and self-reported skill use are in a range that is comparable to prior assessments. Given these results, and the successful link across PIAAC and two prior surveys – IALS and ALL – the PIAAC database can be considered a source for consistent and valid comparisons across countries and subpopulations within countries. Good comparability was achieved over time and across assessment modes.

18.2.2 Classification of items into different proficiency levels

After estimation of the item parameters and respondents' proficiencies (person parameters) in the item calibration stage, items were classified into different proficiency levels separately for each cognitive domain. The purpose of classifying items into different levels is to provide more descriptive information about group proficiencies. That is, the different item levels provide information about the underlying or latent characteristics of an item; the higher the latent characteristic (which reflects our understanding of literacy skills), the higher the level. This item classification into different levels is done by selecting a response probability (RP) value (which defines a point on the scale for that the item function has a certain probability) to predict the probability of correctly responding to a group of items that share characteristics and then to use the selected RP value to assign items to the different proficiency levels. Each level is defined by certain score boundaries for each domain.

While the definitions of the score boundaries for the literacy and numeracy domains are the similar, the score boundaries for PSTRE are different. As there were fewer problem-solving items (14 items) than items from the other domains (2 x 76 items), and the problem-solving items were more difficult, only three levels were defined for this domain. Table 18.7 shows the score boundaries used in PIAAC for literacy and numeracy, and Table 18.8 shows the score boundaries for PSTRE. The decision for the score boundaries was based on expert judgment utilizing the distribution of item difficulties.

[6] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 18.7: Score boundaries for item classification for the domains of literacy and numeracy

Level	Literacy - Score	Numeracy - Score
below level 1	0-175	0-175
1	176-225	176-225
2	226-275	226-275
3	276-325	276-325
4	326-375	326-375
5	376-500	376-500

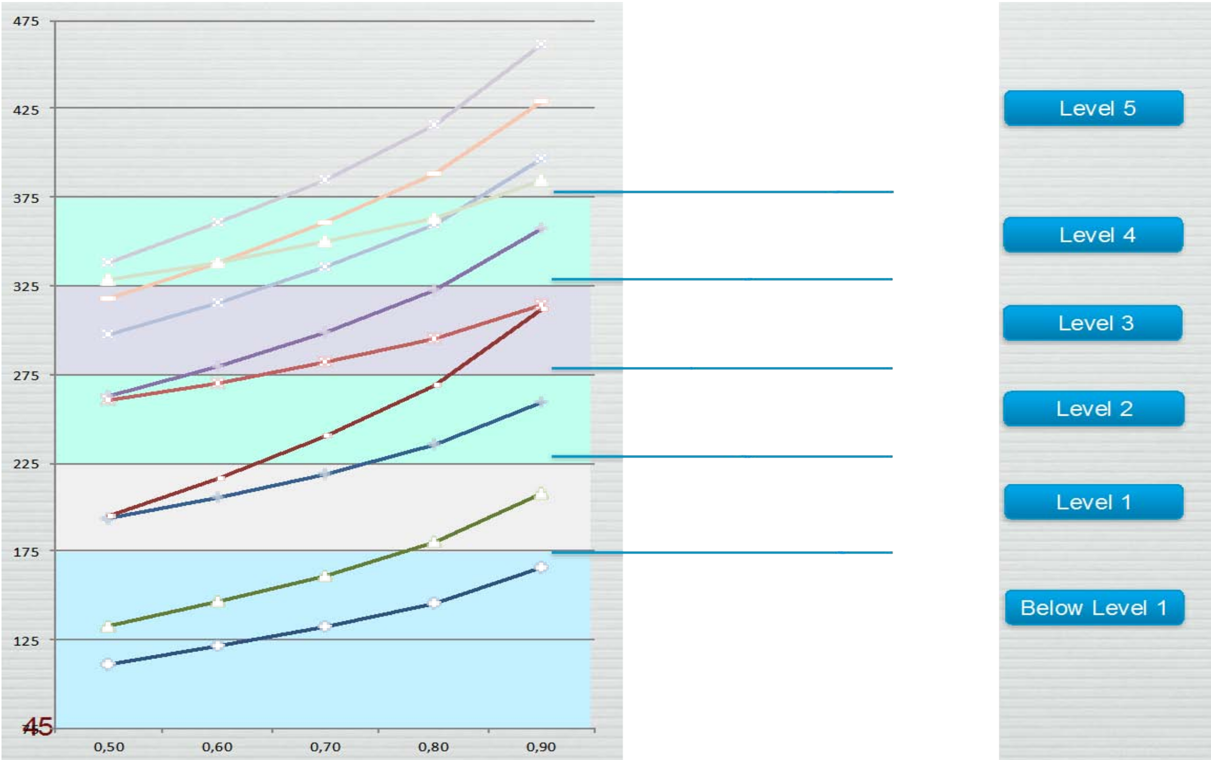
Table 18.8: Score boundaries for item classification for the domain of PSTRE

Level	PSTRE – Score
below level 1	0-240
1	241-290
2	291-340
3	341-500

So far, there is no generally agreed upon rule in the research literature that has been used to characterize items along a proficiency scale. RP values around .65 have been used in most school-based surveys, while values as high as .80 have been used in some adult surveys including IALS and ALL. More recently, however, the US National Academy of Sciences recommended that the National Assessment of Adult Literacy (NAAL) survey (the most recent US survey of adults) use an RP value more closely aligned with school-based surveys. For PIAAC it was decided to use an RP value of .67; some countries received an additional RP value of .80 at their request for the purpose of a better comparability with prior surveys (IALS, ALL). Items are assigned to different proficiency levels due to the selected RP value.

As shown in Figure 18.4, the selection of the RP value impacts where a particular item is classified along the scale. While the selection of an RP value can impact the level in which an item is located, the selection of an RP value has no impact on the proficiency distribution or the percentage of respondents who fall within a particular level (see Figure 18.4).

Figure 18.4: Example for the impact of selected RP values on the placement of items along a scale



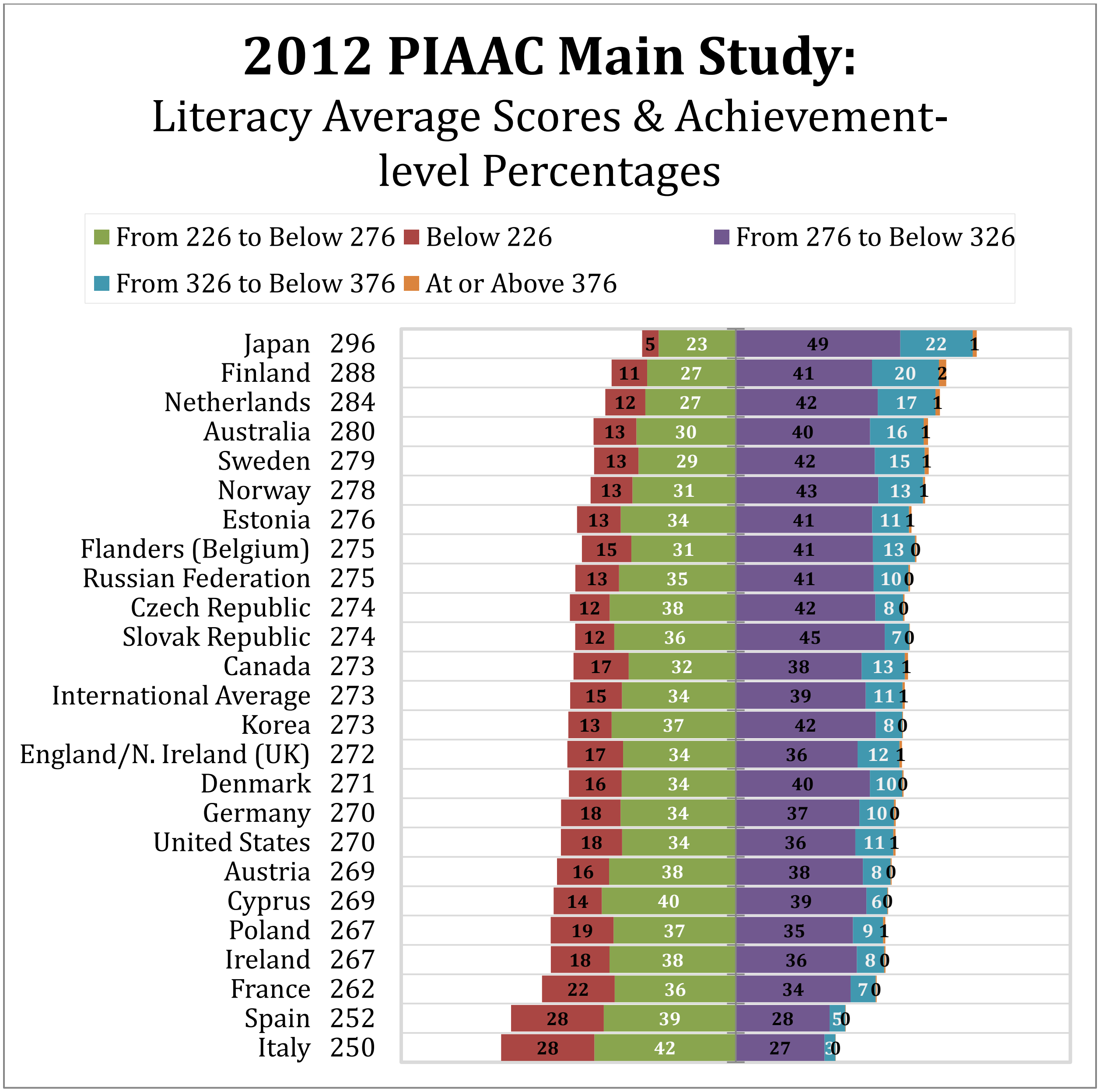
Note: The x-axis in the left hand side exhibit is the response probability (RP) and the y-axis denotes the scale score of the domain.

It is also important to keep in mind that the precision of measurement along a scale is not impacted by the RP value. The same items define the underlying scale regardless of which RP value is selected. Finally, it is important to note that the RP value does not decide on which item measures in which level: All items contribute to the measurement precision in all levels of proficiency, the RP value is one point on the item function graph at which a certain probability is reached. Respondents with a proficiency located below this point have a lower probability (but not 0.0) than the RP value chosen, and respondents with a probability above this point have a higher probability (but do not solve the item with certainty) of solving an item. That means that an item that was located in level 4 using an RP value of 0.67 will also provide information on respondents that are located in levels 3 or 5. The location of an item at a certain level simply implies that (for the chosen RP value) this item is most representative of that particular level.

Chapter 21 describes the content definition for each proficiency level per cognitive domain. Figures 18.5 to 18.7 show the percentage of respondents per country at each level of proficiency for each cognitive domain (note that France and the Russian Federation[7] are not included in the figures as their data were not received in time).

[7] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

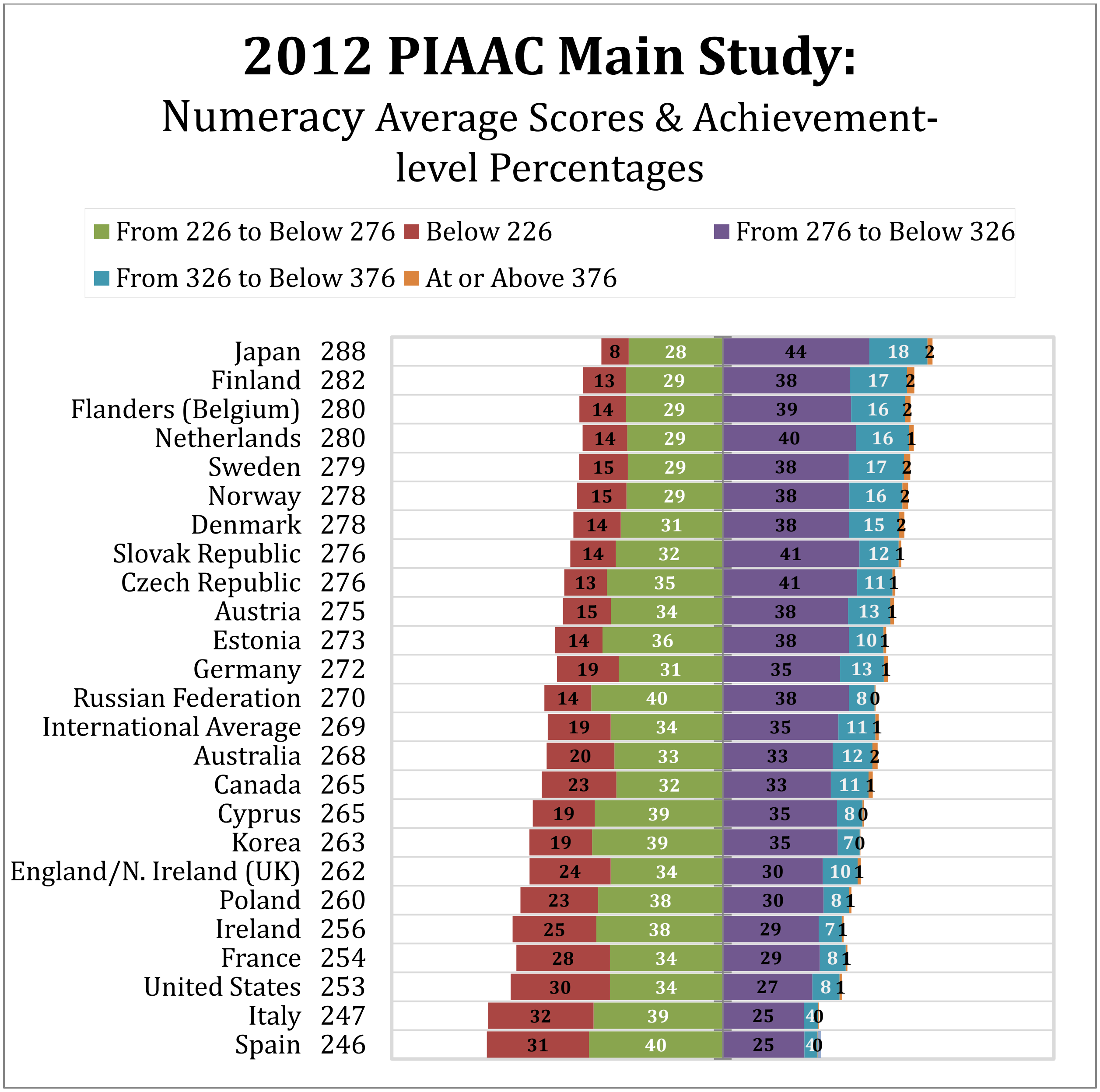
Figure 18.5: Percentage of respondents per country[8] at each level of proficiency for the domain of literacy



---

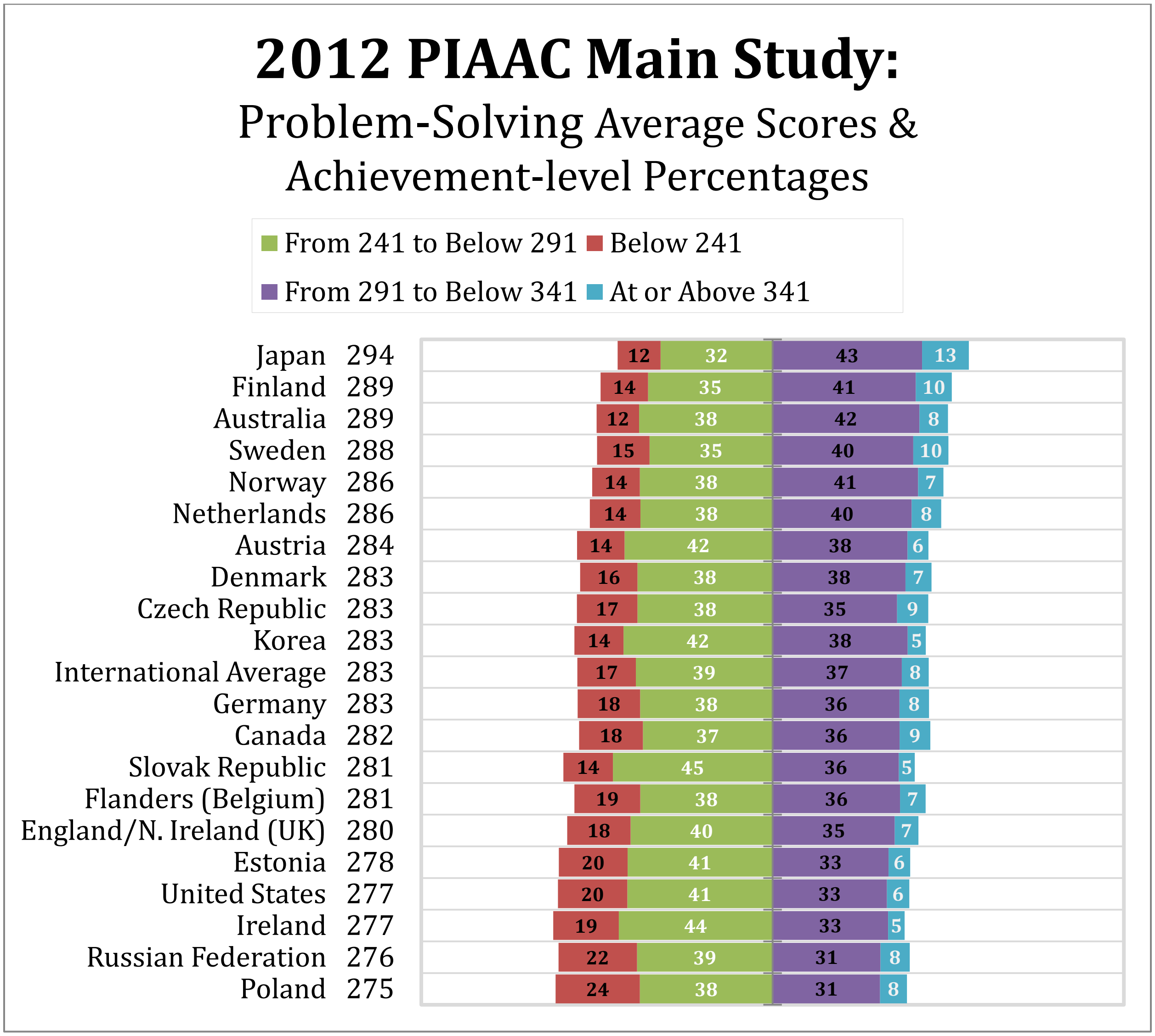
[8] Please refer to the note regarding the Russian Federation, and notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Figure 18.6: Percentage of respondents per country[9] at each level of proficiency for the domain of numeracy



[9] Please refer to the note regarding the Russian Federation, and notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Figure 18.7: Percentage of respondents per country[10] at each level of proficiency for the domain of PSTRE



18.2.3 Transforming the plausible values to PIAAC scales

The plausible values (derived from the population modeling) were transformed using a linear transformation to form a scale that is linked through anchor items to IALS and ALL for literacy and numeracy. This scale can be used to compare the overall performance of countries or subgroups within a country. It can also be used to compare performance along the scale based on statistical criteria such as percentiles.

The linear transformation is based on a concurrent calibration of the literacy and numeracy scales across all countries participating in PIAAC, and also includes data from countries that

[10] Please refer to the note regarding the Russian Federation, and notes A and B regarding Cyprus in the *Note to Readers* section of this report.

participated in IALS and ALL. The reported country distributions from IALS and ALL were used to align the IRT-based country distributions for PIAAC, IALS and ALL to ensure comparability between the three assessments.

To compare the proficiency estimates of the different countries with regard to the cognitive domains, the weighted mean of each of the 10 plausible values per country, and then the average of these 10 means was calculated. Table 18.9 shows the average plausible values for each cognitive domain per country as well as the resampling-based standard errors.

Table 18.9: Average plausible values and resampling-based standard errors per country for the PIAAC domains of literacy, numeracy, and PSTRE

	Literacy		Numeracy		PSTRE	
Country	Average Plausible Values	Standard Error	Average Plausible Values	Standard Error	Average Plausible Values	Standard Error
Australia	280	0.9	268	0.9	289	0.9
Austria	269	0.7	275	0.9	284	0.7
Canada	273	0.6	265	0.7	282	0.7
Cyprus[11]	269	0.8	265	0.8	---	---
Czech Rep.	274	1.0	276	0.9	283	1.1
Denmark	271	0.6	278	0.7	283	0.7
England/N. Ireland (UK)	272	1.0	262	1.1	280	0.9
Estonia	276	0.7	273	0.5	278	1.0
Finland	288	0.7	282	0.7	289	0.8
Flanders (Belgium)	275	0.8	280	0.8	281	0.8
France	262	0.6	254	0.6	---	---
Germany	270	0.9	272	1.0	283	1.0
International Avg. (OECD)	273	0.2	269	0.2	283	0.2
Ireland	267	0.9	256	1.0	277	1.0
Italy	250	1.1	247	1.1	---	---
Japan	296	0.7	288	0.7	294	1.2
Korea	273	0.6	263	0.7	283	0.8
Netherlands	284	0.7	280	0.7	286	0.8
Norway	278	0.6	278	0.8	286	0.6
Poland	267	0.6	260	0.8	275	1.3
Russian Fed.[12]	275	2.7	270	2.7	276	4.3

[11] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[12] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 18.9 (cont.): Average plausible values and resampling-based standard errors per country for the PIAAC domains of literacy, numeracy, and PSTRE

	Literacy		Numeracy		PSTRE	
Country	Average Plausible Values	Standard Error	Average Plausible Values	Standard Error	Average Plausible Values	Standard Error
Slovak Rep.	274	0.6	276	0.8	281	0.8
Spain	252	0.7	246	0.6	---	---
Sweden	279	0.7	279	0.8	288	0.6
United States	270	1.0	253	1.2	277	1.1

18.3 Analysis of data with plausible values

If the scale proficiency values (*θ)* were known for all respondents, it would be possible to directly compute any statistic t(*θ,y*), for example, a scale or composite subpopulation sample mean, a sample percentile point, or a sample regression coefficient to estimate a corresponding population quantity T.

Because the scaling models are latent variable models, *θ* values are not observed. To overcome this problem, we follow the approach taken by Rubin (1987) and treating *θ* as "missing." data The value t(*θ*,*y*) is approximated by its expectation given (*x*,*y*), the data actually observed, as follows:

$$t^{*}(\bar{x},\bar{y}) = E[t(\bar{\theta},\bar{y})|\bar{x},\bar{y}] = \int t(\bar{\theta},\bar{y})p(\bar{\theta}|\bar{x},\bar{y})d\theta$$

It is possible to approximate t^* using plausible values (also referred to as imputations) instead of the unobserved *θ* values. Plausible values are random draws from the conditional distribution of the scale proficiencies given the item responses x_j, background variables y_j, and model parameters (see section 17.2.). For any respondent, the value of *θ* used in the computation of *t* is replaced by a randomly selected value from the respondent's conditional distribution. Rubin (1987) argues that this process should be repeated several times so that the uncertainty associated with imputation can be quantified. For example, the average of multiple estimates of *t*, each computed from a different set of plausible values, is a numerical approximation of t^* in the above equation; the variance among them reflects uncertainty due to not observing *θ*. It should be noted that this variance does not include any variability due to sampling from the population.

It cannot be emphasized too strongly that the plausible values are not a substitute for test scores for individuals. Plausible values incorporate responses to test items and information about the background of responses and can therefore not be used to compare individual test takers in the usual sense. Plausible values are only intermediary computations in the calculation of the integrals in the above equation in order to estimate population characteristics such as subgroup means and standard deviations. When the underlying model is correctly specified, plausible values will provide consistent estimates of population characteristics, even though they are not generally unbiased estimates of the proficiencies of the individuals with whom they are associated (von Davier, Gonzalez & Mislevy, 2009), provide examples and a more detailed explanation). The key idea lies in a contrast between plausible values and the more familiar ability estimates of educational measurement that are in a sense optimal for each respondent (e.g., bias corrected maximum likelihood estimates, which are consistent estimates of a respondent's proficiency θ, and Bayesian estimates, which provide minimum mean-squared errors with respect to a reference population). Point estimates that are optimal for individual respondents have distributions that can produce decidedly nonoptimal (inconsistent) estimates of population characteristics (Little & Rubin, 1983). Plausible values, on the other hand, are constructed explicitly to provide consistent estimates of population effects. For further discussion, see Mislevy, Beaton, Kaplan, and Sheehan (1992).

After obtaining the plausible values from the posteriori distribution, they can be employed to evaluate the previous equation for an arbitrary function T as follows:

1) Using the first vector of plausible values for each respondent, evaluate T as if the plausible values were the true values of θ. Denote the result T_1.

2) In the same manner as in step 1 above, evaluate the sampling variance of T, or Var(T_1,), with respect to respondents' first vectors of plausible values. Denote the result Var_1.

3) Carry out steps 1 and 2 for the second through all 10 vectors of plausible values, thus obtaining T_u and Var_u for u=2, . . .,10.

4) The best estimate of T obtainable from the plausible values is the average of the 10 values obtained from the different sets of plausible values:

   $T. = \frac{\sum_u T_u}{10}$

5) An estimate of the variance of T is the sum of two components: an estimate of Var(T_u) obtained as in step 4 and the variance among the T_us:

$$Var(T.) = \frac{\sum_u Var_u}{10} + (1 + \frac{1}{10}) \frac{\sum_u (T_u - T.)^2}{10 - 1}$$

The first component in *Var(T.)* reflects uncertainty due to sampling from the population; the second component reflects uncertainty because the respondents' proficiencies θ are only indirectly observed through *x* and *y*.

Example for partitioning the estimated error variance:

The following example illustrates the use of plausible values (PV) for partitioning the error variance. Tables 18.10a-c present data for nine subgroups of respondents with differing employment status (variable C_Q07: 1 = full-time employed or self-employed; 2 = part-time employed or self-employed; 3 = unemployed; 4 = pupil or student; 5 = apprentice or internship; 6 = in retirement or early retirement; 7 = permanently disabled; 9 = fulfilling domestic tasks of looking after family; 10 = other). Ten plausible values were calculated for each respondent for each scale (domain). Each column in these tables presents the means of these 10 plausible values.

Table 18.10a: Example for use of plausible values to partitioning the error – PVs 1 to 5

		Plausible Value									
		1		2		3		4		5	
C_Q07	N	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
1	2532	276.14	1.51	276.22	1.59	275.51	1.52	275.82	1.41	275.20	1.57
2	602	267.38	7.18	267.67	6.05	268.15	7.34	265.97	6.85	266.94	5.56
3	414	248.64	6.92	249.27	5.74	249.86	5.59	250.40	7.07	250.87	6.14
4	442	278.88	5.86	279.50	7.00	278.95	7.60	277.38	5.81	279.51	5.36
5	14	261.22	115.05	278.57	75.31	277.95	75.11	266.04	137.08	273.69	128.94
6	203	266.33	13.80	266.51	13.62	268.66	12.41	271.01	12.60	266.97	12.87
7	270	229.81	8.12	231.01	8.32	228.63	10.57	229.45	8.47	230.05	7.54
9	281	269.96	10.06	267.22	13.44	268.92	11.81	270.63	10.01	269.02	11.24
10	137	272.87	29.97	273.99	26.47	269.86	38.14	273.93	32.09	270.52	30.41

Table 18.10b: Example for use of plausible values to partitioning the error – PVs 6 to 10

		Plausible Value									
		6		7		8		9		10	
C_Q07	N	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
1	2532	275.74	1.65	275.60	1.50	275.66	1.58	274.70	1.53	275.65	1.53
2	602	269.03	5.00	266.45	6.58	267.41	6.25	268.85	6.85	266.45	6.38
3	414	250.63	6.21	249.98	5.78	249.53	7.05	248.78	6.27	251.82	6.97
4	442	279.61	5.84	278.27	6.78	279.04	6.07	282.19	5.37	279.11	5.62
5	14	284.81	46.85	272.05	162.29	296.01	59.46	267.64	159.43	280.77	71.99
6	203	267.92	17.15	268.15	18.38	265.38	13.60	268.05	17.40	267.07	14.06
7	270	230.91	9.76	228.51	9.89	229.83	8.29	230.72	9.81	230.06	8.73
9	281	268.73	13.24	266.79	11.60	268.63	14.09	270.38	11.23	269.29	15.18
10	137	272.85	32.07	270.49	34.29	273.31	31.97	275.00	29.46	272.68	35.37

Table 18.10c: Example for use of plausible values to partitioning the error – sample error, measurement error, and standard error based on the 10 PVs

C_Q07	N	Mean of 10 PVs	Sampling Error	Measurement Error	Standard Error
1	2532	275.62	1.24	0.46	1.32
2	602	267.43	2.53	1.08	2.75
3	414	249.98	2.52	1.03	2.73
4	442	279.24	2.48	1.29	2.79
5	14	275.88	10.16	10.60	14.68
6	203	267.61	3.82	1.63	4.15
7	270	229.90	2.99	0.91	3.13
9	281	268.96	3.49	1.30	3.73
10	137	272.55	5.66	1.79	5.94

The error variance, or squared standard error, of the mean plausible values differs greatly for the subgroups. The error variance reflects a component of error with regard to the lack of precision of the measurement instrument and a component of error with regard to sampling. The variance can be reduced by either increasing the precision of the measurement instrument (for example, increasing the number of items) or increasing the sample size. The resampling method was used to estimate the variance due to sampling using the each set of imputed values. This component of variance is similar across the 10 plausible values; the size is influenced by the homogeneity of proficiencies among respondents in a subgroup but not by the sample size or by the precision of the survey instruments. The sampling error is smaller when the subgroup consists of respondents with similar proficiencies. The total error variance can be calculated as the summation of "sampling error" and measurement error."

The last column presents the standard error of the subpopulation mean, which is equal to the square root of the sum of the two components' variance. Pairwise differences can be evaluated using these standard errors. However, multiple comparisons such as the six possible pairwise comparisons of this example need to consider an adjustment of significance level such as Hochberg Stagewise Procedure (HSP), described in Hochberg (1988).

Hochberg developed a method for multiple comparisons that utilizes the order of significance levels among all comparisons. HSP begins by placing the comparisons in an increasing order of significance levels, i.e., $P_1 \leq P_2 \leq \ldots \leq P_3 \leq \ldots P_M$. It proceeds to sequentially evaluate P_j with adjusted critical significance level of $\alpha/(m-j+1)$ where α is the target significance level. If P_j is smaller than the critical significance level then the process continues until a non-significance comparison is found. All preceding comparisons before the first nonsignificant comparison are declared significant and all subsequent comparisons are declared nonsignificant. Both the Bonferoni method and the HSP control the Type 1 error of false discovery of significant comparison when in fact it is nonsignificant. The False Discovery Rate (FDR) procedure (Benjamini & Hochberg, 1995) controls the expected proportion of falsely rejected hypotheses, finding the comparison nonsignificant when in fact it is significant. The procedure is very similar to HSP for ordering the comparisons by the significance level, then using the critical significance level of α*j/m for j-th in the comparisons. The determination of the significance of comparisons is identical to the HSP.

The standard errors of mean proficiencies, percentages, and percentiles play an important role in interpreting subpopulation results and in comparing the performances of two or more subpopulations. The resampling standard errors reported by PIAAC are statistics whose quality depends on certain features of the samples from which the estimates are obtained. In certain cases, primarily when the standard error is based on a small number of respondents, the mean squared error associated with the estimated standard errors may be quite large.

18.4 Developing common scales across modes of administration and for the purpose of trends

As described in section 17.4, the linking design for PIAAC aims to link items and booklets across different assessment modes as well as to the IALS and ALL surveys to provide trend measures. PIAAC items were linked between PBA and CBA and to items from IALS and ALL. Common scales were obtained through item calibration using an IRT analysis (2PL, GPCM) and a linear transformation of the new estimates using the group means and standard deviations obtained from the previous IRT estimates in IALS and ALL (see section 17.4.2). The comparability of item parameters across countries, modes of assessment, and time (different surveys) were evaluated. If a large deviation from the common item parameters was observed for one or more countries, unique item parameters were estimated for the deviant item.

The following section present the results of this linking design for the PIAAC Main Study.

18.4.1 Linking outcomes

The linking design for the PIAAC Main Study was aimed at establishing comparability across countries with regard to both PBA and CBAs as well as the link between PIAAC and the IALS and ALL surveys. This pertains especially to the paper- and computer-based items in numeracy, and the paper-based items in literacy; deviations were limited to a few items and countries. The PIAAC item parameters for a few computer-based literacy items (which were adapted from IALS and ALL paper-based items) were not comparable with the item parameters of IALS and ALL, and with item parameters of the paper-based PIAAC assessment. By estimating new item parameters – that is, parameters were estimated for the CBA only – for those computer-based literacy items, comparability improved to the level of numeracy. The majority of linking items shared the common item parameters, that is, parameters were estimated for the data of the PBA and the CBA together.

The proportion of respondents who received the 12 different adaptive paths for the literacy scale varied between 5.0 to 13.5% across countries. For the numeracy scale, the proportions varied between 2.9% to 16.7% among paths and countries. The following two tables (18.11, 18.12) present the distribution of the 12 routing paths for literacy and numeracy scales by country, showing that the distributions are comparable between countries. A note on notation: L13 means that literacy testlets 1 and 3 were administered for the stage 1 and 2.

Table 18.11: Distribution of routing paths for the literacy module by country

Country	CBA Core	Literacy Routing Path											
		L11	L12	L13	L14	L21	L22	L23	L24	L31	L32	L33	L34
Australia	0.01	0.06	0.08	0.08	0.08	0.06	0.08	0.08	0.09	0.06	0.11	0.09	0.12
Austria	0.01	0.08	0.08	0.08	0.07	0.07	0.08	0.08	0.09	0.07	0.11	0.09	0.10
Canada	0.01	0.07	0.08	0.08	0.08	0.07	0.09	0.08	0.09	0.06	0.10	0.08	0.12
Cyprus[13]	0.01	0.07	0.09	0.08	0.07	0.07	0.08	0.08	0.08	0.09	0.10	0.09	0.08
Czech Rep.	0.01	0.07	0.09	0.08	0.08	0.07	0.08	0.09	0.08	0.07	0.09	0.08	0.10
Denmark	0.01	0.07	0.08	0.08	0.08	0.07	0.09	0.08	0.09	0.07	0.10	0.08	0.11
England/N. Ireland (UK)	0.02	0.08	0.09	0.08	0.09	0.07	0.09	0.08	0.08	0.06	0.09	0.08	0.11
Estonia	0.01	0.07	0.07	0.09	0.08	0.07	0.08	0.08	0.09	0.07	0.09	0.09	0.11
Finland	0.01	0.06	0.07	0.08	0.09	0.07	0.09	0.07	0.10	0.06	0.10	0.09	0.11
Flanders (Belgium)	0.02	0.06	0.07	0.09	0.09	0.06	0.08	0.09	0.09	0.06	0.10	0.08	0.11
France	0.02	0.07	0.07	0.09	0.08	0.07	0.09	0.08	0.08	0.06	0.10	0.08	0.10
Germany	0.01	0.07	0.08	0.09	0.08	0.07	0.08	0.07	0.08	0.07	0.09	0.09	0.11
Ireland	0.01	0.07	0.08	0.07	0.08	0.06	0.09	0.08	0.10	0.06	0.10	0.09	0.11
Italy	0.01	0.09	0.08	0.08	0.08	0.08	0.09	0.08	0.07	0.07	0.10	0.08	0.09
Japan	0.00	0.06	0.08	0.07	0.09	0.05	0.08	0.08	0.10	0.06	0.09	0.10	0.14
Korea	0.01	0.06	0.08	0.07	0.09	0.06	0.09	0.08	0.08	0.08	0.10	0.10	0.11
Netherlands	0.01	0.07	0.08	0.08	0.08	0.07	0.09	0.08	0.09	0.06	0.10	0.09	0.11
Norway	0.02	0.06	0.07	0.08	0.09	0.07	0.09	0.09	0.09	0.05	0.09	0.09	0.11
Poland	0.01	0.07	0.08	0.08	0.09	0.07	0.09	0.08	0.09	0.06	0.09	0.08	0.10
Russian Fed.[14]	0.00	0.07	0.07	0.07	0.09	0.06	0.10	0.08	0.11	0.06	0.10	0.09	0.10
Slovak Rep.	0.01	0.07	0.08	0.08	0.08	0.07	0.08	0.08	0.09	0.09	0.09	0.09	0.09
Spain	0.01	0.08	0.09	0.07	0.07	0.08	0.08	0.09	0.08	0.08	0.09	0.09	0.09
Sweden	0.01	0.08	0.07	0.08	0.08	0.07	0.09	0.09	0.09	0.06	0.09	0.09	0.11
United States	0.02	0.07	0.08	0.08	0.08	0.06	0.08	0.09	0.08	0.06	0.09	0.08	0.12

[13] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[14] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 18.12: Distribution of routing paths for the numeracy module by country

Country	CBA core	Numeracy Routing Path											
		N11	N12	N13	N14	N21	N22	N23	N24	N31	N32	N33	N34
Australia	0.01	0.06	0.08	0.07	0.08	0.06	0.09	0.09	0.11	0.04	0.09	0.09	0.13
Austria	0.01	0.06	0.07	0.07	0.09	0.05	0.08	0.08	0.11	0.05	0.10	0.09	0.13
Canada	0.01	0.07	0.08	0.08	0.09	0.06	0.08	0.08	0.10	0.05	0.09	0.09	0.13
Cyprus[15]	0.01	0.06	0.09	0.08	0.08	0.05	0.08	0.08	0.09	0.05	0.10	0.10	0.13
Czech Rep.	0.01	0.07	0.08	0.08	0.09	0.05	0.09	0.09	0.10	0.04	0.08	0.10	0.12
Denmark	0.01	0.06	0.08	0.07	0.09	0.05	0.08	0.08	0.11	0.05	0.10	0.10	0.13
England/N. Ireland (UK)	0.02	0.07	0.09	0.09	0.06	0.07	0.09	0.08	0.09	0.05	0.08	0.08	0.13
Estonia	0.01	0.06	0.08	0.08	0.09	0.05	0.08	0.09	0.11	0.04	0.09	0.10	0.12
Finland	0.01	0.07	0.08	0.09	0.09	0.05	0.08	0.09	0.10	0.04	0.08	0.09	0.13
Flanders (Belgium)	0.02	0.06	0.08	0.07	0.09	0.05	0.08	0.08	0.10	0.04	0.10	0.09	0.14
France	0.02	0.07	0.07	0.08	0.08	0.06	0.08	0.08	0.10	0.05	0.09	0.09	0.12
Germany	0.01	0.07	0.08	0.08	0.09	0.06	0.08	0.07	0.11	0.04	0.09	0.09	0.13
Ireland	0.01	0.07	0.08	0.08	0.08	0.06	0.09	0.08	0.09	0.06	0.10	0.10	0.12
Italy	0.01	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.11	0.05	0.09	0.08	0.10
Japan	0.00	0.06	0.07	0.08	0.09	0.04	0.08	0.09	0.11	0.03	0.10	0.09	0.17
Korea	0.01	0.06	0.08	0.08	0.07	0.05	0.08	0.09	0.11	0.05	0.10	0.09	0.14
Netherlands	0.01	0.07	0.08	0.08	0.09	0.05	0.08	0.09	0.10	0.05	0.08	0.09	0.13
Norway	0.02	0.07	0.08	0.08	0.09	0.05	0.08	0.08	0.09	0.04	0.09	0.10	0.13
Poland	0.01	0.06	0.07	0.09	0.09	0.05	0.08	0.08	0.10	0.05	0.09	0.09	0.13
Russian Fed.[16]	0.00	0.08	0.09	0.08	0.09	0.07	0.08	0.09	0.08	0.05	0.09	0.08	0.12
Slovak Rep.	0.01	0.06	0.07	0.08	0.09	0.06	0.09	0.08	0.10	0.05	0.09	0.08	0.13
Spain	0.01	0.07	0.09	0.09	0.07	0.07	0.09	0.09	0.09	0.05	0.09	0.09	0.11
Sweden	0.01	0.07	0.08	0.08	0.10	0.05	0.08	0.08	0.10	0.04	0.08	0.10	0.12
United States	0.02	0.07	0.08	0.07	0.07	0.06	0.09	0.09	0.08	0.05	0.09	0.09	0.13

[15] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[16] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *Series B (Methodological) 57*, 289–300.

Chen, H., Yamamoto, K., & von Davier, M. (in press). Controlling MST exposure rates in international large-scale assessments. In D. Yan, D., A. A. von Davier, & C. Lewes (Eds.), *Computerized Multistage Testing –Theory and Applications*. Taylor & Francis.

Efron, B. (1982). The jackknife, the bootstrap, and other resampling plans. *Society of Industrial and Applied Mathematics CBMS-NSF Monographs*, 38.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, *75*, 800-802.

Little, R. J. A. & Rubin, D. B. (1983). On jointly estimating parameters and missing data. *American Statistician, 37, 218-220.*

Mislevy, R. J., Beaton, A., Kaplan, B. A., & Sheehan, K. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement, 29,* 133-161.

Mislevy, R. J. & Sheehan, K. (1987) Marginal estimation procedures. In A. E. Beaton, *Implementing the new design: The NAEP 1983-84 technical report (pp. 293-360). (No. 15-TR-20)* Princeton, NJ: Educational Testing Service

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: John Wiley & Sons.

von Davier, M., Gonzalez, E., & Mislevy, R. (2009). What are plausible values and why are they useful? In: *IERI Monograph Series: Issues and methodologies in Large Scale Assessments, Vol. 2*. Retrieved from IERI website: http://www.ieriinstitute.org/IERI_Monograph_Volume_02_Chapter_01.pdf

von Davier, M. Sinharay, S., Oranje, A., & Beaton, A. (2006). Statistical Procedures used in the National Assessment of Educational Progress (NAEP): Recent Developments and Future Directions. In C.R. Rao and S. Sinharay (Eds.), *Handbook of Statistics (Vol. 26): Psychometrics.* Amsterdam: Elsevier.

Appendix 18.1: Marginal correlations (Pearson) per country of the cognitive domains literacy (LIT), numeracy (NUM) and problem solving in technology rich environments (PSL), respectively, with scales of the BQ, based on the 10 plausible values obtained from the population modeling (conditioning)

	Learning at work	Readiness to learn	Use of ICT skills at home	Use of ICT skills at work	Use of influencing skills at work	Use of num skills at home	Use of num skills at work	Use of planning skills at work	Use of reading skills at home	Use of reading skills at work	Use of task discretion at work	Use of writing skills at home	Use of writing skills at work
Australia													
LIT	0.010	0.323	0.313	0.214	0.183	0.313	0.151	0.116	0.335	0.211	0.146	0.28	0.210
NUM	0.003	0.284	0.277	0.188	0.169	0.319	0.200	0.118	0.310	0.207	0.127	0.253	0.195
PSTRE	0.015	0.204	0.307	0.218	0.086	0.217	0.121	0.045	0.122	0.103	0.094	0.171	0.133
Austria													
LIT	0.073	0.267	0.315	0.259	0.180	0.266	0.230	0.108	0.327	0.290	0.095	0.257	0.208
NUM	0.055	0.250	0.267	0.226	0.197	0.276	0.260	0.116	0.311	0.288	0.120	0.222	0.194
PSTRE	0.095	0.175	0.350	0.255	0.080	0.241	0.190	0.022	0.235	0.167	0.024	0.218	0.130
Canada													
LIT	-0.016	0.271	0.291	0.195	0.168	0.264	0.154	0.107	0.337	0.214	0.177	0.229	0.175
NUM	-0.031	0.237	0.253	0.175	0.139	0.273	0.194	0.098	0.296	0.193	0.158	0.193	0.148
PSTRE	-0.007	0.191	0.318	0.190	0.093	0.249	0.114	0.041	0.220	0.108	0.161	0.230	0.098
Cyprus[17]													
LIT	0.000	0.108	0.080	0.047	0.040	0.075	0.062	-0.023	0.168	0.099	-0.006	0.094	0.121
NUM	0.034	0.146	0.078	0.107	0.083	0.144	0.146	0.018	0.209	0.180	0.026	0.141	0.159
PSTRE													
Czech Rep.													
LIT	0.056	0.212	0.248	0.131	0.189	0.255	0.162	0.112	0.324	0.196	0.079	0.195	0.172
NUM	0.063	0.188	0.202	0.158	0.200	0.230	0.200	0.136	0.303	0.240	0.118	0.140	0.164
PSTRE	0.075	0.210	0.295	0.178	0.158	0.307	0.217	0.095	0.288	0.154	0.106	0.215	0.159

[17] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Appendix 18.1: Marginal correlations (Pearson) per country of the cognitive domains literacy (LIT), numeracy (NUM) and problem solving in technology rich environments (PSL), respectively, with scales of the BQ, based on the 10 plausible values obtained from the population modeling (conditioning)

	Learning at work	Readiness to learn	Use of ICT skills at home	Use of ICT skills at work	Use of influencing skills at work	Use of num skills at home	Use of num skills at work	Use of planning skills at work	Use of reading skills at home	Use of reading skills at work	Use of task discretion at work	Use of writing skills at home	Use of writing skills at work
Denmark													
LIT	0.057	0.209	0.307	0.228	0.184	0.269	0.180	0.067	0.328	0.22	0.066	0.255	0.152
NUM	0.030	0.155	0.246	0.233	0.181	0.251	0.235	0.087	0.284	0.231	0.074	0.193	0.141
PSTRE	0.076	0.151	0.323	0.224	0.099	0.291	0.203	-0.009	0.212	0.119	0.004	0.263	0.110
England/N. Ireland (UK)													
LIT	0.041	0.277	0.291	0.218	0.207	0.250	0.187	0.145	0.314	0.271	0.161	0.253	0.209
NUM	0.018	0.280	0.269	0.187	0.175	0.269	0.205	0.133	0.302	0.261	0.152	0.221	0.189
PSTRE	0.049	0.275	0.368	0.289	0.171	0.262	0.226	0.098	0.254	0.222	0.152	0.269	0.195
Estonia													
LIT	0.016	0.277	0.278	0.217	0.125	0.273	0.169	0.056	0.315	0.197	0.163	0.229	0.155
NUM	0.026	0.274	0.246	0.217	0.150	0.270	0.219	0.096	0.313	0.220	0.184	0.204	0.161
PSTRE	0.042	0.276	0.403	0.24	0.130	0.322	0.184	0.045	0.255	0.183	0.187	0.270	0.143
Finland													
LIT	-0.014	0.176	0.322	0.219	0.182	0.303	0.167	0.102	0.313	0.207	0.066	0.279	0.171
NUM	-0.020	0.133	0.280	0.235	0.131	0.315	0.253	0.071	0.283	0.204	0.081	0.220	0.157
PSTRE	0.036	0.148	0.384	0.216	0.090	0.315	0.152	0.039	0.236	0.084	0.057	0.265	0.074
Flanders (Belgium)													
LIT	0.100	0.272	0.329	0.274	0.219	0.267	0.24	0.103	0.343	0.322	0.129	0.238	0.218
NUM	0.082	0.261	0.299	0.235	0.230	0.265	0.269	0.130	0.320	0.316	0.142	0.219	0.221
PSTRE	0.136	0.192	0.376	0.288	0.146	0.306	0.226	0.092	0.234	0.243	0.101	0.202	0.188

Appendix 18.1: Marginal correlations (Pearson) per country of the cognitive domains literacy (LIT), numeracy (NUM) and problem solving in technology rich environments (PSL), respectively, with scales of the BQ, based on the 10 plausible values obtained from the population modeling (conditioning)

	Learning at work	Readiness to learn	Use of ICT skills at home	Use of ICT skills at work	Use of influencing skills at work	Use of num skills at home	Use of num skills at work	Use of planning skills at work	Use of reading skills at home	Use of reading skills at work	Use of task discretion at work	Use of writing skills at home	Use of writing skills at work
France													
LIT	0.175	0.274	0.318	0.210	0.222	0.332	0.230	0.113	0.395	0.344	0.146	0.272	0.230
NUM	0.154	0.268	0.297	0.226	0.232	0.338	0.279	0.136	0.392	0.376	0.148	0.257	0.243
Germany													
LIT	0.036	0.256	0.335	0.217	0.179	0.327	0.207	0.076	0.368	0.262	0.101	0.230	0.158
NUM	0.031	0.247	0.291	0.214	0.196	0.329	0.252	0.088	0.339	0.250	0.123	0.201	0.148
PSTRE	0.033	0.172	0.373	0.235	0.087	0.331	0.211	0.002	0.265	0.157	0.035	0.224	0.120
Ireland													
LIT	0.044	0.242	0.262	0.194	0.148	0.226	0.169	0.129	0.320	0.243	0.136	0.245	0.192
NUM	0.050	0.225	0.236	0.176	0.133	0.243	0.215	0.106	0.306	0.242	0.160	0.227	0.168
PSTRE	0.018	0.155	0.355	0.273	0.075	0.229	0.181	0.059	0.232	0.144	0.144	0.239	0.131
Italy													
LIT	0.030	0.220	0.241	0.133	0.20	0.214	0.204	0.110	0.362	0.281	0.085	0.167	0.210
NUM	0.043	0.219	0.238	0.145	0.193	0.239	0.258	0.113	0.371	0.263	0.123	0.159	0.213
Japan													
LIT	0.073	0.243	0.208	0.174	0.09	0.143	0.170	0.001	0.271	0.153	0.015	0.052	0.095
NUM	0.041	0.250	0.194	0.216	0.161	0.179	0.256	0.048	0.279	0.205	0.078	0.061	0.119
PSTRE	0.023	0.178	0.267	0.246	0.03	0.151	0.195	-0.036	0.149	0.088	0.011	0.023	0.064
Korea													
LIT	0.092	0.327	0.304	0.183	0.200	0.317	0.220	0.080	0.370	0.252	0.035	0.153	0.163
NUM	0.073	0.300	0.271	0.154	0.197	0.298	0.211	0.073	0.351	0.223	0.033	0.132	0.149
PSTRE	0.092	0.164	0.309	0.208	0.045	0.203	0.119	-0.051	0.129	0.063	-0.049	0.125	0.108

Appendix 18.1: Marginal correlations (Pearson) per country of the cognitive domains literacy (LIT), numeracy (NUM) and problem solving in technology rich environments (PSL), respectively, with scales of the BQ, based on the 10 plausible values obtained from the population modeling (conditioning)

	Learning at work	Readiness to learn	Use of ICT skills at home	Use of ICT skills at work	Use of influencing skills at work	Use of num skills at home	Use of num skills at work	Use of planning skills at work	Use of reading skills at home	Use of reading skills at work	Use of task discretion at work	Use of writing skills at home	Use of writing skills at work
Netherlands													
LIT	0.055	0.339	0.372	0.241	0.185	0.278	0.178	0.102	0.336	0.252	0.175	0.280	0.188
NUM	0.050	0.301	0.330	0.218	0.185	0.278	0.229	0.110	0.320	0.235	0.170	0.249	0.170
PSTRE	0.071	0.274	0.365	0.208	0.152	0.304	0.158	0.049	0.276	0.157	0.141	0.240	0.142
Norway													
LIT	0.043	0.203	0.279	0.273	0.212	0.238	0.183	0.141	0.283	0.236	0.128	0.176	0.178
NUM	-0.011	0.150	0.235	0.282	0.175	0.236	0.227	0.138	0.243	0.227	0.116	0.142	0.169
PSTRE	0.066	0.184	0.335	0.257	0.128	0.295	0.224	0.048	0.243	0.131	0.050	0.197	0.134
Poland													
LIT	0.039	0.253	0.301	0.182	0.151	0.306	0.206	0.092	0.384	0.226	0.080	0.264	0.130
NUM	0.016	0.235	0.267	0.158	0.133	0.314	0.223	0.101	0.338	0.199	0.077	0.222	0.136
PSTRE	0.004	0.126	0.317	0.149	0.072	0.255	0.124	0.011	0.260	0.107	0.077	0.207	0.064
Russian Federation[18]													
LIT	0.052	0.180	0.177	0.080	0.014	0.225	0.088	0.032	0.273	0.098	0.070	0.084	0.088
NUM	0.110	0.180	0.160	0.054	0.048	0.249	0.119	0.019	0.259	0.122	0.073	0.066	0.103
PSTRE	0.092	0.194	0.263	0.122	0.049	0.225	0.139	0.057	0.279	0.166	0.072	0.112	0.140
Slovak Rep.													
LIT	0.120	0.293	0.121	0.087	0.098	0.256	0.125	0.061	0.370	0.193	0.082	0.135	0.139
NUM	0.133	0.319	0.169	0.121	0.115	0.278	0.173	0.101	0.388	0.226	0.119	0.165	0.156
PSTRE	0.007	0.131	0.188	0.15	0.055	0.131	0.125	0.036	0.151	0.095	0.081	0.089	0.129

[18] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Appendix 18.1: Marginal correlations (Pearson) per country of the cognitive domains literacy (LIT), numeracy (NUM) and problem solving in technology rich environments (PSL), respectively, with scales of the BQ, based on the 10 plausible values obtained from the population modeling (conditioning)

	Learning at work	Readiness to learn	Use of ICT skills at home	Use of ICT skills at work	Use of influencing skills at work	Use of num skills at home	Use of num skills at work	Use of planning skills at work	Use of reading skills at home	Use of reading skills at work	Use of task discretion at work	Use of writing skills at home	Use of writing skills at work
Spain													
LIT	0.067	0.227	0.300	0.190	0.212	0.305	0.180	0.138	0.393	0.273	0.074	0.261	0.224
NUM	0.054	0.227	0.252	0.185	0.207	0.291	0.215	0.143	0.375	0.268	0.072	0.242	0.234
Sweden													
LIT	-0.013	0.201	0.276	0.253	0.171	0.219	0.207	0.038	0.279	0.184	0.042	0.203	0.156
NUM	0.014	0.178	0.222	0.227	0.16	0.223	0.243	0.063	0.244	0.181	0.065	0.152	0.127
PSTRE	0.070	0.204	0.377	0.269	0.107	0.276	0.253	-0.001	0.266	0.111	-0.018	0.264	0.114
United States													
LIT	-0.064	0.205	0.267	0.183	0.119	0.240	0.111	0.116	0.245	0.184	0.174	0.20	0.137
NUM	-0.036	0.179	0.245	0.199	0.118	0.235	0.141	0.112	0.220	0.186	0.152	0.164	0.131
PSTRE	-0.041	0.147	0.333	0.22	0.081	0.222	0.105	0.08	0.138	0.099	0.158	0.177	0.126

Note: The correlations for the ICT scales might be underestimated as not every respondent received the ICT items according to the path of the adaptive testing.

Chapter 19: Proficiency Scale Construction

Kentaro Yamamoto, Lale Khorramdel and Matthias von Davier, ETS

19.1 Overview

In this chapter we describe and illustrate the development of scales and items (based on respective frameworks) for the cognitive part of the PIAAC survey as well as the evaluation of the items and the instrument through a field test.

The Field Test addressed three main areas: a) operational (in terms of feasibility of implementation), b) instrumentation, and c) scaling and psychometric characteristics, and was important for the successful implementation of the Main Study. The fact that the results of PIAAC had to be linked to previous assessments, while also being implemented in both PBA and CBA modes (including an adaptive aspect), added to that importance.

Results of the Field Test provided information and guidance with regard to the sampling, data collection, refinement of scoring procedures for the CBA items, inference strategies, and analysis methods for the Main Study.

19.2 Development of the described scales

In the following, we will refer to the term “task” as an umbrella term for “item” as well as “item group associated with a common stem.” A task can have a more complex structure compared to an item representing the construct or scale of interest, while an item is a question referring to a common stem or stimulus. Thus, one task can have one or multiple items. In the context of the description of the frameworks and scale developments, we refer to “tasks”; in the context of data analyses, we refer to “items.”

19.2.1 Stage 1: Identifying possible scales

The identification and definition of scales (domains) to be measured in international large-scale assessments are important as they provide a foundation for the design of the assessment and set the boundaries for what will be included. PIAAC 2012 assessed the three main domains of literacy, numeracy, and problem solving and thus has to give definitions for them. All three domains are multifaceted constructs referring to complex competencies. The following section provides an overview of these definitions, explains on which prior definitions and assessments they are based, and explains to which extent prior definitions were expanded to meet new opportunities and changes in society.

Literacy scale:

The definition of the *literacy* scale in PIAAC is based on the previous adult literacy assessments known as IALS and ALL, but extends these assessments because adults have faced new literacy opportunities since (e.g., the use of email and other digital media) those assessments were created. Therefore, it was necessary to broaden the literacy construct to include new modes of text. PIAAC also provides an opportunity to deepen our understanding of the cognitive skills that underlie adult literacy and the role that engagement plays in literacy. While in IALS and ALL the literacy scale was divided into the scales *prose literacy* (continuous texts) and *document literacy* (noncontinuous texts), PIAAC joins them into one *literacy* scale. On the one hand, the concept of literacy in PIAAC was defined to support a link to the IALS and ALL assessments to enable the analysis of trends. On the other, it was expanded in three ways:

1) The range of texts to be considered should be broader than in previous assessments; in particular, the definition should include those texts often identified as electronic texts.
2) The type of cognitive activities identified should go beyond simply using text, to enable a deeper understanding of literacy ability.
3) The concept of literacy should also include engagement in literacy practices.

The Literacy Expert Group defines the PIAAC literacy scale as follows: "*Literacy is understanding, evaluating, using and engaging with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential.*"

Excursus:

The following definitions and explanations provide a deeper understanding of the literacy definition that is used for PIAAC 2012:

- *Written text:* PIAAC aims to expand the range of texts which were assessed by IALS and ALL (informative texts of both continuous and noncontinuous form) to include a greater variety of text types, such as narrative and interactive texts, and a greater variety of media (computer, PDA, Blackberry or iPhone, etc.). Including electronic text opens the assessment to new types of text and content. Some of these novel form/content combinations include interactive texts, such as exchanges in comments sections of blogs or in email response threads, multiple texts, whether displayed at the same time on a screen or linked through hypertext, and expandable texts, where a summary can be linked to more detailed information if the user chooses.

- *Understanding:* Understanding means the construction of meaning (large and small, literal and implicit) from text. This can be as basic as understanding the meaning of the words, or as complex as comprehending the underlying theme of a lengthy argument or narrative. PIAAC aims to provide a more direct measure of understanding (not just an indirect one). While the assessment of reading components provides the construct to support basic understanding, the assessment of literacy in PIAAC includes tasks that explicitly tap more complex understanding, such as the relation(s) between different parts of the text, the gist of the text as a whole, and insight into the author's intent. Readers also have to understand the social function of each text and the way this influences structure and content.

- *Evaluating:* Readers continually (have to) make judgments about a text and evaluate information in terms of accuracy, reliability and timeliness. This is particularly important with online material as, in contrast to published print information, online information is more varied, ranging from authoritative sources to postings with unknown or uncertain authenticity.

- *Using:* Using means that the reader approaches the text with a specific task in mind, that is, reading is directed toward applying the information and ideas in a text to an immediate task or to reinforce or change beliefs. In some cases, using a text in this way requires just minimal understanding – getting the meaning of the words with some elementary recognition of structure. In others, it requires using both syntactic and more complex structural understanding to extract the information.

- *Engaging with:* Adults differ in how engaged they are with reading texts and how much a role reading plays in their lives (reading because it is required versus reading for pleasure). Studies have found that engagement with reading is an important correlate with the direct cognitive measures.

- *Participate in society:* Adults use text as a way to engage with their social surroundings, to learn about and to actively contribute to life in their community, close to home and more broadly. For many adults, literacy is essential to their participation in the labor force. Thus, literacy has a social aspect. It is a part of the interactions between and among individuals.

- *Achieve one's goals:* Literacy is increasingly complicit in meeting those needs, whether simply finding one's way through shopping, or negotiating complex bureaucracies whose rules are commonly available only in written texts. It is also important in meeting adult needs for sociability, entertainment and leisure, and work.

- *Develop one's potential:* Surveys suggest that many adults engage in some kind of learning throughout their life, much of it self-directed and informal. Much of this learning requires some use of text, and as individuals want to improve their life, whether at work or outside, they need to understand, use, and engage with printed and electronic materials.

In PIAAC texts are organized in three ways:

1) *Medium (print and digital):* A major development of PIAAC over previous adult surveys is the inclusion of digital (or electronic) texts. Because some texts that are applied electronically are just simple copies of printed texts, digital texts are not distinguished by the medium in which they occur, but by whether they make use of text navigation and display features found only through digital devices. Any text that could appear on a printed page exactly as it appears on a screen is considered a *print* text; any text that could not appear on a printed page with all its features intact is considered a *digital* text.

2) *Format (continuous and noncontinuous):* In IALS and ALL, texts were classified as

continuous (prose literacy) or noncontinuous (document literacy). This is an important distinction, as each requires different text knowledge and a different approach to text processing. At the same time, many actual texts involve some elements that are continuous and some that are noncontinuous. Thus, the distinction is better made on the basis of what type(s) of text a task requires.

a. Continuous: This type of text is conventionally made up of sentences formed into paragraphs. Some continuous texts include typographic features, such as indenting and headings, that signal the organization of the text, but many do not. Examples of continuous texts include newspaper and magazine articles, brochures, manuals, emails, and many web pages.

b. Noncontinuous: This type of text uses explicit typographic features, rather than paragraphs, to organize information. While there may be full sentences in some noncontinuous texts, most consist of words or phrases organized by some kind of matrix arrangement. Tables, graphs, charts and forms are all examples of noncontinuous texts.

c. Combined: This type of text has both continuous and noncontinuous elements. Examples of mixed texts include web pages with a list of links, newspaper articles that incorporate line graphs or pie charts, and brochures with attached order forms.

d. Multiple: Multiple texts consist of texts that have been generated and which make sense independent of each other. The texts are juxtaposed or loosely linked for a particular purpose. The relationships among the component texts need not be obvious. The texts may be contradictory or complementary. Such texts are common in digital settings, but are also found in print environments.

3) *Type (rhetorical stance of the text):* The IALS and ALL frameworks are classified as continuous texts by their rhetorical stance, because all share the same structure, but noncontinuous texts also share the same rhetorical stances. Therefore, in PIAAC, the stances of all types of text were identified using the six categories employed in the IALS and ALL assessments (the text type "hypertext" was eliminated in PIAAC because it is not a rhetorical category but a structural type which will be included under electronic text for PIAAC). The point of having rhetorical stance as a variable is not due to evidence that difficulty is affected by it, but as a way of ensuring that a variety of texts are included on the assessment. The six types of rhetorical stance for PIAAC are as follows:

a. Description: This is the type of text where the information refers to properties of objects in space. A page of a manual that identifies the parts of some device, such as a Cuisinart, is a description, as is a verbal depiction of a piece of art.

b. Narration: This is the type of text where the information refers to properties of objects in time. Stories recounted to make a point, such as fables, are narrations, as are texts about the steps an individual took to solve a problem.

c. Exposition: In this type of text, information is presented as composite concepts or mental constructs, or those elements into which concepts or mental constructs can be analyzed. The text provides an explanation of how the component elements interrelate in a meaningful whole. A text that explains the nature of some health problem or one that tells about the election process in the United States would be an exposition.

d. Argumentation: This type of text presents propositions as to the relationship among concepts or other propositions. An important subclassification of argument texts is persuasive texts. Newspaper editorials are one example, as are advertisements.

e. Instruction (sometimes called injunction): This type of text provides directions on what to do. Most equipment manuals contain instruction texts, but so do other guides, such as those about first aid or a leisure activity.

f. Records: Records are texts that are designed to standardize, present and conserve information without embedding in other stances. A table of standings in a sports league is an example of a record, as is a graph of the changes in oil prices. The minutes of a meeting constitute another type of record.

More detailed information about how to classify noncontinuous texts (Matrix Documents, Graphic Documents, Locative Documents, Entry Documents, Combination Documents) and electronic texts (Hypertext, Index-like, Interactive) is given in the PIAAC literacy framework (OECD, 2012).

Because both the motivation to read and the interpretation of the content may be influenced by the *context*, a fair assessment must include material from a broad range of settings in order to include some material that would be familiar to any participant. PIAAC tried to include the following contexts (or content areas):

- Home and family
- Health and safety
- Community and citizenship
- Consumer economics
- Work
- Leisure and recreation
- Education and training

Furthermore, the following three *cognitive operations with text* can be identified that are needed when working on items or tasks:

1) Access and identify information in the text
2) Integrate and interpret (relate parts of text to each other)
3) Evaluate and reflect (understanding of the text as a whole)

As a supplement to the main literacy assessment, PIAAC includes an additional assessment of *reading components*. This assessment aims to provide information on the reading abilities of adults with poor skills in order to get a proper understanding of their difficulties. The following five reading components were identified:

- Alphanumeric perceptual knowledge and familiarity
- Word recognition
- Word knowledge (vocabulary)
- Sentence processing
- Passage fluency

More detailed information about contexts, cognitive operations, and further points that influence the difficulty of items (such as the transparency of items, semantic complexity, amount of information needed, prominence of information, and competing information), as well as more information about the reading components, is given in the PIAAC literacy framework (OECD, 2012).

Numeracy scale:

Basic computational or mathematical knowledge has always been considered part of the fundamental skills that adults need to function well and be able to accomplish various goals in their everyday, work and social life. Societies now present increasing amounts and wider ranges of information of a quantitative nature to citizens from all walks of life in diverse contexts. As workplaces are becoming more concerned with involving all workers in improving efficiency and quality, the importance of numeracy skills is growing. Numeracy involves, among other things, the handling of arithmetical processes, understanding of proportions and probabilistic ideas, understanding of numerical, geometric and graphical types, and representations of quantitative information, critical interpretation of statistical or mathematical messages, and ability to solve various types of quantitative problems.

The Numeracy Expert Group defines the PIAAC numeracy scale as follows: "*Numeracy is the ability to access, use, apply, interpret, and communicate mathematical information and ideas in order to effectively manage and respond to the mathematical demands of diverse situations in the information age.*"

The conceptualization of numeracy is based on the previous adult literacy assessments IALS and ALL, as well as on a review of scholarly literature and research findings (with regard to IALS, the numeracy scale in PIAAC is most closely related to the scales of document literacy and quantitative literacy). Numeracy operates on two levels:

> It relates to numeracy as a construct describing a competence as defined above, and to numerate behavior, which is the way a person's numeracy is manifested in the face of situations or contexts, which have mathematical elements or carry information of a quantitative nature. In this way, inferences about a person's numeracy are possible through analysis of performance on assessment tasks designed to elicit numerate behavior.
>
> In congruence to the view of numeracy as a competence, numeracy will be described as comprising both cognitive elements (i.e., various knowledge bases and skills) as well as noncognitive or semicognitive elements (i.e., attitudes, beliefs, habits of mind, and other dispositions) which together shape a person's numerate behavior.

The Numeracy Expert Group gives the following definition for numerate behavior: "Numerate behavior involves managing a situation or solving a problem in a real context, by responding to mathematical content/information/ideas represented in multiple ways."

Thus, numerate behavior comprises four facets:

a) contexts (everyday life, work, societal, further learning)
b) responses (identify, locate or access; act upon, use; order, count, estimate, compute, measure, model; interpret; evaluate/solve; communicate)
c) mathematical content/information/ideas (quantity and number; dimension and shape; pattern, relationships, change; data and chance)
d) representations of mathematical information (objects and pictures; numbers and mathematical symbols; formulae; diagrams and maps, graphs, tables; texts; technology-based displays)

A more detailed definition of these four facets is given in the PIAAC numeracy framework (OECD, 2012).

Numeracy is required so people can effectively cope with or respond to a range of situations that are embedded in the course of life with real, personal meaning to them. Three key types of situations are given below to illustrate the range of numeracy demands placed on adults:

- *Generative situations*: These demand that people count, quantify, compute, or otherwise manipulate numbers, concrete objects, visual elements, and so forth, to create/generate new numbers or estimates (e.g., calculating the total price of products while shopping, finding the number of boxes in a crate, measuring the area of a room to be painted in order to calculate the amount of materials needed to do the job, reading a menu and computing the cost of a specified meal, filling out an order form for a product, figuring out travel times between train stations based on a timetable, etc.). The numerical information in many types of generative situations may be evident in the situation itself (e.g., real objects to be arranged, sorted, counted, or measured; a graph on a computer display) or may also be communicated through text or embedded in different types of text; hence, such situations may also involve language skills to varying degrees.

- *Interpretive situations*: These demand that people make sense, and grasp the implications of, messages that contain information of a mathematical or statistical nature but that *do not involve direct manipulation of numbers* (e.g., deciding whether a generalization stated in a newspaper article about results from a recent opinion poll is valid; other examples can be added where references to proportions, averages, samples, bias, correlation, risk, or causality are discussed or implied, such as in the context of genetic or medical counseling, or understanding of statistical process control displays).

- *Decision situations*: These demand that people locate and consider multiple pieces of information in order to determine a course of action, typically in the presence of conflicting goals, constraints or uncertainty. Two key subtypes here are *optimization tasks* (identification of optimal ways to use resources such as money or supplies, or schedule personnel or time) and *choice tasks* (making choices among alternatives, such as which of several apartments to rent, which pension or health insurance plan to join, whether to undergo a surgical medical procedure that has known probabilities of certain side effects, etc.). It is important to note that optimization and choice tasks can be part of a broader problem-solving process, where alternatives have to be generated and then evaluated. Thus, what is being termed here a decision situation at times also can be viewed as a problem-solving situation.

The three types of numeracy situations described above are not mutually exclusive, and other cases may exist, possibly of a hybrid nature. Moreover, it is important to keep in mind the impact of evolving technologies (Internet- or technology-based resources).

While it is possible to define numeracy in general terms without invoking literacy, the structure of the tasks and demands in adults' lives show these areas cannot be considered mutually exclusive. Mathematical or statistical information is carried by or embedded in text in some, but certainly not all, contexts in which adults have to function. To the extent this happens, one's performance on numeracy tasks will depend not only on formal mathematical or statistical knowledge but possibly also on literacy related factors such as vocabulary, reading comprehension, reading strategies, or prior literacy experiences.

Problem solving scale:

The aim of PIAAC to assess problem solving in technology-rich environments (PS-TRE) was based on the fact that digital technologies have deeply transformed the way individuals learn, communicate, work, and, more generally, the way they function in societies. Microcomputers, laptops, mobile phones, and the Internet have provided users with powerful tools to search for and make use of immense repertories of information and services. Increasingly versatile mobile technologies allow users to stay connected almost regardless of where they are and what they are doing. And the integration of digital tools in homes, cars and appliances potentially increases the safety, flexibility, and effectiveness of many activities of everyday life.

Yet using computers or other digital devices to perform personal or work-related activities and to solve problems often presents a challenge for the everyday user. People often have trouble installing, setting up, and learning how to use new digital devices and software applications. Users often confine themselves to a few basic, but ineffective, procedures. Then, even routine computer use for mundane tasks is often prone to errors, delays and incidents. Tools and technologies are normally meant to facilitate the resolution of a problem. They may, however, also contribute to making a problem more difficult, especially when a person has limited knowledge and experience with those tools and technologies.

Therefore, PIAAC aimed to analyze the problem-solving skills involved in the uses of digital technologies, thus concentrating on problems people deal with when using ICT. Those problems share the following characteristics:

- The existence of the problem is primarily a *consequence of the availability of new technologies*. One example relates to the vast amount of information now available on the Web. This gives rise to problems related to locating and evaluating information for quality and credibility, for example, when seeking advice about legal issues or medical conditions. Other examples include the increasing capacity of electronic storage devices, with the subsequent problems of organizing and sorting large numbers of files; or the growing practice of social communication on the Web, with the subsequent problem of learning and making use of new social norms as far as private vs. public information.

- The problem solution *requires the use of computer-based artifacts* (tools, representational formats, computational procedures) that were not available previously, or at least not available to the general public. An example is the management of personal finance by using spreadsheets, statistical packages, and graphical tools. Here the problem itself may not be new (i.e., keeping spending in balance with income), but the new artifacts modify the distribution of work across social agents (professional vs. laypersons) and deeply transform the procedures and steps required to solve the problem.

- The problems are *related to the handling and maintenance of technology-rich environments* themselves (e.g., how to operate a computer, how to fix a settings problem, how to use the Internet browser in a technical sense).

Understanding and evaluating meaningful information available in technology-rich environments is central to the construct of problem solving. Most of the problems require one to handle vast amounts of symbolic information and, thus, the ability to deal with semantic content or meaning (e.g,. understanding command names in dropdown menus, naming of files and folders, hits in a search engine, or links in a Web page). Furthermore, many problems require the person to read and understand electronic texts, graphics and numerical data.

The Problem Solving Expert Group defines the PIAAC problem-solving scale as follows: "*Problem solving in technology-rich environments (PS-TRE) involves using digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks. PIAAC 2012 will focused on the abilities to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, accessing and making use of information through computers and computer networks.*"

More information and specific comments on the words and phrases used in this definition is given in the PIAAC problem-solving framework (OECD, 2012).

The PIAAC domain of problem solving may be organized along three key dimensions:

Cognitive dimensions: the mental structures and processes by which a person actually performs problem solving (goal setting and monitoring progress; planning; locating and evaluating information; and selecting, organizing, and transforming information)
Technologies: the devices, applications and functionalities through which problem solving is conducted (hardware devices; simulated software applications; commands and functions; representations such as text and graphics, etc.)
Tasks or problem statements: elements of a situation that trigger a condition for problem solving (scenario and task directions presented to test takers; specific material conditions in which the test is organized)

More detailed information and examples of the different key dimensions of the PIAAC problem-solving scale are given in the PIAAC problem-solving framework (OECD, 2012).

Even if the domains of literacy, numeracy, and problem solving rely on the same "core" cognitive processes (e.g., the ability to decode printed symbols, working memory capacity), there are aspects that distinguish problem solving from the other two domains:

- As problem solving specifically assesses goal setting, monitoring, and planning in technology-rich environments, problem-solving tasks emphasize the processes of problem finding and problem shaping that are typical of problem solving. Problem-solving tasks also focus on the kinds of problems that are associated with these environments (e.g., problems associated with Web-based texts that are not well defined and the need for logical operators to search for information).

- Problem-solving tasks were carried out in environments that involve multiple, complex sources of information. Some of the tasks even required the test taker to use multiple environments and to shift across them. Thus, problem solving assessed decision making as far as information sources to be used (e.g., the act of choosing which environment to use or whether or not to go to another website). Evaluation was included as a critical underlying part of problem solving. Additionally, selecting appropriate devices or tools took a more prominent role for this domain.

- In terms of information processing, problem solving is a specific construct in that: a) it focuses on the pragmatic evaluation of sources in terms of reliability and the adequacy of information relative to the problem statement as opposed to mere topical relevance, which is more applicable for literacy; b) it focuses on the integration of information across sources, especially in cases where the sources provide inconsistent information.

19.2.2 Stage 2: Design principles and constraints (selecting items for the assessment)

During the item development process and the assignment of items per PIAAC domain (scale) to the assessment, the following principles were taken into consideration:

a) *Items should cover as many aspects as possible with regard to the different text types, contexts and processes of literacy, the different facets and contexts of numeracy, and the different cognitive dimensions and contexts of problem solving.* Items should require the activation of a broad range of skills and knowledge included in these constructs, as portrayed in the conceptual frameworks.

b) *Items should aspire to maximal authenticity and cultural appropriateness.* Items should be derived from real-life stimuli and pertain to situations that can be expected to be of importance or relevant in different contexts in at least some of the countries participating in PIAAC. Item content and questions should appear purposeful to respondents across cultures, even if they are not necessarily familiar to all adults in all countries.

c) *Items should have a free-response format, to the extent feasible by the computer platform used for administering the direct assessments in PIAAC.* Items should be structured to include a stimulus (e.g., a picture, drawing, visual display) and one or more questions, the answers to which the respondent communicates via the modes available on the computer, primarily: entry, click, highlight a region of the stimulus, usage of various pull-down menus. (Text entry is limited to very specific words or sometimes a simple number due to the concerns listed above regarding the inability to score text entries with keying/typing errors, and the presence of multiple ways to express the same content in words and/or numbers).

d) *Items should spread over different levels of ability.* Items should span the range of ability levels anticipated within PIAAC participants, from low-skilled individuals (which are of interest in countries where policies and educational programs may be earmarked for low-skill populations) all the way to those with advanced competencies. The need to reduce the number of items to be administered in any one domain has led to the practice (in previous assessments as well as in PIAAC) of including few very easy items (i.e., items at level 1) and few very hard items (i.e., items at Level 5). Respondents will be classified at Level 1 if they could not do well on Level 2 tasks. Likewise, those classified at Level 5 will be those who performed well on Level 4 items and on the few real Level 5 items.

   It follows that a more detailed assessment of the specific skills of Level 1 respondents requires a separate diagnostic assessment. Therefore, the reading components assessment was conducted in PIAAC. To enable the adaptive testing process and thus reach an efficient estimation of respondents' ability levels, the following distribution of items at the different difficulty levels was sought for constructing the item pool for literacy and numeracy (there was no adaptive testing for problem solving in technology-rich environments) for the main PIAAC assessment, based on the results of the Field Test (pilot test) in 2010: 5% Level 1 items, 25% Level 2 items, 40% Level 3 items, 25% Level 4 items, and 5% Level 5 items.

e) *Items should vary in the degree to which the task is embedded in text.* Some items should be embedded in or include relatively rich texts, while others should use little or no text. This distribution aimed to reflect the different levels of text involvement in real-world numeracy tasks, as well as reduce overlap with the literacy scale.

f) *Items should be efficient.* To allow for coverage of many key facets of the literacy, numeracy, and problem-solving competencies, the inclusion of a large number of diverse stimuli and questions was needed. However, in light of testing time constraints, the use of short items was necessitated, precluding items that could simulate extended problem-solving processes or require a lengthy open-ended response.

g) *Items should be adaptable to unit systems across participating countries.* Items should be designed so that their underlying literacy/mathematical/problem solving demands are as consistent as possible across countries regarding language and conventions. For example, items were designed so that different currency systems or different systems of measurement (metric or imperial) could be applied to the numbers or figures used. Items should retain equivalency with respect to their literacy/mathematical/problem solving or cognitive demands after being translated.

In addition to the above listed principles, the assignment of items to the PIAAC assessment design had to address two further points: the linking between PIAAC and previous surveys, and the link between the CBA and PBA. To enable the linking among PIAAC, IALS and ALL, a part of the PIAAC item pool came from the IALS and ALL surveys (approximately 60%), while the other part consisted of new items that were developed for PIAAC. With regard to the literacy scale, the newly developed PIAAC items had to be assigned either to the subscale "prose literacy" or the subscale "document literacy" as the scale "literacy" was divided into these subscales in IALS and ALL. To enable the link between the PBA and the new CBA, a portion of the IALS and ALL items, which were all paper-based, had to be redesignated to be administered within the CBA. Furthermore, a portion of the newly developed items had to be assigned to both modes of assessment. Altogether, a larger portion of the IALS and ALL items as well as the newly developed PIAAC items was used for the CBA, while a smaller portion was used for the PBA. The latter procedure had not only the aim of enabling the linking design but also to provide a reliable and valid assessment for adults who were unfamiliar or uncomfortable with computers.

Due to the limited testing time (only 60-70 minutes for the core part, the cognitive adaptive assessment, and the BQ), it was decided to use a larger number of short tasks for the scales of literacy and numeracy (in order to cover all relevant contexts and facets) instead of a smaller number of more complex tasks, although it is recognized that ability to solve complex or extended literacy and numeracy problems is an inherent part of these competencies.

PIAAC also aimed to include open-ended response formats, with the limitation that the computer system (TAO) in the current stage of development could not accept most types of free-form text-based answers because of the huge possible diversity in how respondents may enter answers. The limitations stem from the difficulty of automatically coding the responses in dozens of languages while accommodating various grammatical and syntactical structures, as well as overcoming typing mistakes, which are naturally expected when people type text into a computer. Some workarounds were implemented to capture selected types of open-ended responses and circumvent the text-processing limitation to some extent, for example, by using multiple pull-down menus that allow a respondent to "construct" a response from predesigned elements or response ranges. Maybe in future cycles of PIAAC, some of the current technical limitations will be resolved, allowing for better coverage of more aspects of the assessed constructs.

19.2.3 Stage 3: Field Test – Aims, design, and data collection

After developing new items (for literacy and numeracy) and new measures (for reading components and problem solving in technology-rich environments), and assigning old and new items to the cognitive domains based on their respective frameworks, the quality of the developed instrument had to be tested and evaluated. More precisely, the scaling and psychometric characteristics of the items had to be evaluated before using the items for the PIAAC Main Study in 2012. Furthermore, it was necessary to evaluate if the linking design was working and providing reliable trend measures, and if the computer delivery platform (for the CBA) was stable and reliable. Thus, a Field Test trial was designed and data analyzed in 2010 to yield adequate information relating to these questions. Moreover, standardized procedures and quality mechanisms were tested in the Field Test; they were embedded into various phases of PIAAC including survey development, implementation, and analysis and reporting of the data. The outcomes of the Field Test were used to assemble the final instruments that were used in the Main Study, and operational issues were modified and refined based on the Field Test. In summary the following areas were evaluated:

- Evaluation of survey operations procedures (data collection procedures, response rates for various subpopulations, data processing including scoring, recoding, and data transmission)
- Quality of the instrument: scaling and psychometric characteristics
- Equivalence of assessment modes: CBA vs. PBA
- Comparability of results between countries
- Trend measure: link between IALS, ALL and PIAAC

The PIAAC Field Test was designed to measures the domains of literacy and reading components, numeracy, and problem solving in technology-rich environments across two modes of administration (paper and pencil and computer delivered), while also offering participating countries both core and optional components. As mentioned earlier, 60 percent of the literacy and numeracy items came from the ALL and the IALS surveys to allow a link to these assessments and provide trend measures.

The full Field Test design assumed 40-45 minutes of administration time for the BQ and JRA and 60 minutes for the direct assessment. The design was based on the sample yield of 1,500 respondents per country/per language (i.e., completed cases) between the ages of 16 to 65: 1,100 for the CBA and 400 for the PBA (with a later modification of a maximum of 200 ICT-core failed samples to be routed to the PBA). On average across 23 countries, 209 respondents failed the ICT-core items, 1,426 completed the BQ section, 830 completed the CBA, and 505 completed the PBA.

Equivalence of scoring standard across countries

Achieving the goal of comparability depends on the equivalence of scoring within and between countries. Scoring was required to determine whether respondents have correctly answered the questions in the paper-based cognitive instruments. Rescoring was conducted as a quality assurance measure to determine whether the scoring rubrics have been applied consistently by every scorer within the country and without bias across the countries.

During the Field Test, participating countries checked the consistency of scoring by having a second scorer rescore 100% of the instruments. Additionally, item-level reliability was conducted to identify items that the scorers had difficulty in scoring consistently. Items with low interrater reliability have been further examined for possible ways to improve scoring accuracy through improved translation, instruction, and/or training for the Main Study.

a) Inter-country scoring reliability (equivalence of scoring of anchor booklets)

In order to evaluate scoring standard across countries, anchor booklets were produced in English (60 anchor booklets for the core part, 60 for literacy, and 60 for numeracy in both the Field Test and the Main Study). This common set of booklets was prepared by test developers and distributed to all countries. Item responses in these booklets were based on actual responses collected in the field as well as responses that reflected key points on which scorers were trained. Because responses were provided in English, scoring teams in each country designated two bilingual scorers responsible for the double-scoring process. Countries were required to follow a specified design to ensure that each booklet was scored twice and that each scorer functioned both as first and second scorer across all the booklets. Scoring results of both scorers were evaluated by the Consortium for consistency between the scorers as well as accuracy against the master scores as designed.

The unit of analysis implemented to evaluate agreement was the number of items multiplied by the number of countries, i.e., (38x23=)836 for the literacy scale and (35x23=)770 for numeracy. Average percentage agreement over items within a country averaged across all countries was 95.7% for literacy items and 95.6% for numeracy items. The variance of average agreements was 2.42 for literacy and 0.06 for numeracy. The number of item by country pairs showing less than 85% agreement was 24 for literacy and 14 for numeracy. Out of those lower agreements, two items were responsible for 12 of 24 for literacy items, and also two items accounted for eight of 14 lower agreements for numeracy. Regarding disagreements per country, there were two countries with more disagreements than the rest of countries. These two countries accounted for 15 of the total of 38 lower agreement item*country pairs.

Altogether, the rescoring of anchor booklets indicated very clearly that scoring of printed cognitive items is accurate, consistent, and without evidence of bias.

b) Intra-country scoring reliability

While reliable scores of anchor booklets ensure comparability of scoring standard across countries, reliability of scoring within a country indicates how accurately such a scoring standard was applied consistently among multiple scorers within a country. Countries followed rescoring instructions that were provided for three-, four- and five-scorer situations.

The unit of analysis implemented to evaluate agreement is identical to one used for the anchor booklets rescoring. Average percentage agreement over items within a country averaged across all countries was 96.1% for literacy items and 97.3% for numeracy items. The variance of average agreements was 10.35 for literacy and 4.00 for numeracy. Number of item by country pairs with less than 85% agreement was 43 for literacy and 13 for numeracy. Out of those lower agreements, three items were responsible for 18 of 43 for literacy items, and one item accounted for eight of 13 lower agreements for numeracy. In terms of country, there were three countries with more disagreements than the rest of the countries. These three countries accounted for 32 out of 52 lower agreement item by country pairs.

Altogether, a small number of items and countries have shown some difficulty in attaining high score reliability. As a consequence, some recommendations were given to optimize the scorer training for the PIAAC Main Study as well as the data capture, operational issues, data transmission, and quality assurance mechanisms.

Instrumentation:

The Field Test addressed the following issues related to instrumentation:

- The accuracy and comparability of survey instruments were reviewed, including translation and scoring guides and all related manuals. These activities resulted in a number of corrections and clarifications.

- The timing and flow of questions in the BQ was evaluated. (The researchers in GESIS performed this task, resulting in the reports included in the summary of BQ instruments.)

- The appropriateness of questions across participating countries was evaluated.

- The response distribution in all categories of the BQ was examined.

The timing information from the Field Test was used to make sure that the Main Study wouldn’t be too long. The Field Test showed that the majority of respondents needed one hour to complete the assessment, and that they were much faster in completing the reading components than expected. Therefore, more items could be included for the Main Study from the existing item pool (one more reading component passage was used in the Main Study than originally planned).

Computer delivery platform:

To evaluate the CBA in PIAAC, the Field Test was delivered on a laptop computer to respondents in their homes. A computer-delivery platform (TAO) integrated with the CAPI tool was used for the administration of the BQ, the JRA and the cognitive instruments. The Field Test addressed the following issues related to the computer-delivery platform:

- The functioning of the cognitive portion of the delivery platform was tested and evaluated (emphasizing response capturing and automatic scoring).

- The functioning of the CAPI system was tested and evaluated (emphasizing the flow of questions and efficiency of the system in capturing information).

- The accuracy of the interviewer's instructions was evaluated.
- The effectiveness of the system during the interview was tested.
- The integration of the PIAAC platform with national survey management systems was verified.

The Field Test showed that no major architectural changes were necessary for the platform, but some system freeze occurred during the test administration that had to be fixed. After addressing this issue in later updates, the Main Study instruments became very stable.

Scaling, psychometric characteristics, equivalence between modes and assessments:

The Field Test data were used to examine scaling methodologies in order to determine the psychometric characteristics of items and scales. This included the evaluation of the equivalence of item parameter estimates among linking items from IALS and ALL to PIAAC, and the equivalence of the estimates between the PBAs and the CBAs. To identify deviations of item-by-country interactions, two measures of mean deviation (MD) and root mean squared deviations (RMSD) were used (see section 17.3.2. for detailed information about the MD and RMSD).

Furthermore, the Field Test was also an opportunity to examine the role of computer familiarity and determine the standards for branching respondents with regard to the adaptive test design of the Main Study. The Field Test provided initial IRT item parameter estimates that were used to construct the adaptive testing algorithm, which was implemented in the Main Study. Thus, the Field Test had to address the following issues with respect to IRT scaling and psychometric characteristics:

- Literacy items were re-estimated using the entire aggregate data of IALS/ALL because the literacy scale is a joint scale of prose and document literacy scales. These new parameter estimates were used for the subsequent analyses. The numeracy scale was introduced in ALL, and subsequent analyses used ALL numeracy estimates.
- In order to examine equivalence of item characteristics across countries, a common set of item parameter estimates of the two-parameter logistic (2PL) model and the general partial credit model (GPCM) was estimated and found to fit quite well to all countries, for all three scales, and in both PBA and CBA. Deviation was fairly small and almost all countries and items were found to be conforming to the international parameter estimates. The sample size in the Field Test was too small for each country to estimate country-specific item parameters.
- Equivalence of item characteristics among the literacy and numeracy items common to IALS and ALL on the paper-and-pencil version was examined. Equivalence of IALS/ALL item parameter estimates to the CBA items adapted from IALS and ALL were also evaluated. Previously estimated IALS and ALL item parameters on PBA fit very well to the PBA items adapted for both scales of literacy and numeracy. For the IALS/ALL items adapted for the PIAAC CBA, previously estimated item parameters fit quite well for the numeracy scales with a few items showing noticeable deviation from the IALS/ALL estimates. For the literacy scales, more items showed clear deviation from

the IALS/ALL estimates. Equivalence of item characteristics of literacy and numeracy items common to PBA and CBA was examined. Several items were freed to estimate CBA only item parameters, while the majority of linking items shared common item parameters between PBA and CBA items.

- Items among the literacy, numeracy and problem-solving items were identified to be assembled into the core assessment.

- The expected proportions of subsamples routed to the different assessment modes and the different stages of the CBA based on preliminary background information and the core were examined. As working with various countries with various ability distributions makes it critical to have a sufficient number of responses for every item, simulation studies were calculated to evaluate item exposure under adaptive procedure.

- The overall psychometric characteristics and quality of the Field Test items were evaluated to guide the selection of items for the Main Study.

The Field Test design

The PIAAC Field Test design provided good item level information on the full range of direct assessment measures and was useful in addressing the other operational and psychometric issues identified above. The BQ and a core set of questions focusing on information and ICT was designed to ensure that respondents who have no familiarity with computers are routed to the PBA. Because the number of respondents without ICT skills could have been numerous, a limitation on the maximum number of respondents was placed at 200 so that the CBA item parameter estimation would not be jeopardized. The limit of 200 respondents was placed to avoid such a scenario. However, most of the countries never reached this limit during the data collection in the Field Test. In order to link the PBA and the CBA, the remaining adults (the majority of adults in each country who are expected to pass the core) were randomly assigned to either one of them. The Field Test design (see Figure 19.1) comprised the following steps and procedures:

Step 1, BQ: The BQ was designed to take 30-40 minutes, and was delivered by the interviewer using a computer-assisted format with respondents taking one of three variable sections (a 20-minute core set of items and one of three, 10-minute subsets that would be administered along with the cognitive instruments). Compared to the original design for the Main Study, the BQ required some modifications to accommodate the large number of questions that go beyond 30-40 minutes (implemented by rotating some of the questions). Moreover, not every respondent answered every question because appropriate questions were presented based on the answer to the previous question(s).

Step 2a, PBA: The PBA was designed to comprise a 10-minute core of either literacy or numeracy skills (each with six items), followed by two 20-minute blocks of literacy or numeracy (totaling 29 items), and a final 10-minute cluster of reading components. Thus, the total testing time was estimated to be 60 minutes. Four paper booklets with varied (balanced) block orders were constructed to control for possible order effects (see Figure 19.1). In the Field Test (as well as the Main Study), every respondent in the PBA took the reading components (in case the international option to assess reading components was chosen by the respective country; see below). But while there was no link between the CBA and the reading components in the Field Test (only respondents working on the PBA also worked on the reading components), respondents who performed poorly on the core and literacy items of the CBA based survey were transferred to the reading components as well.

Step 2b, CBA: The CBA was designed to included twenty-one 60-minute booklets consisting of two 30-minute blocks of items in each booklet. While the items of the CBA in the Main Study were administered adaptively, this was not the case for the Field Test: The block order was balanced, but the item order within each block was fixed. As reflected in this design (see Figure 19.1), each of the computer-delivered booklets contained literacy-only tasks, numeracy-only tasks, literacy and problem-solving tasks, numeracy and problem-solving tasks, or problem solving-only tasks. Overall, for the Field Test, there were thirteen 30-minute blocks that were grouped to form the 21 booklets: four blocks of literacy tasks, four blocks of numeracy tasks, and five blocks of problem-solving tasks.

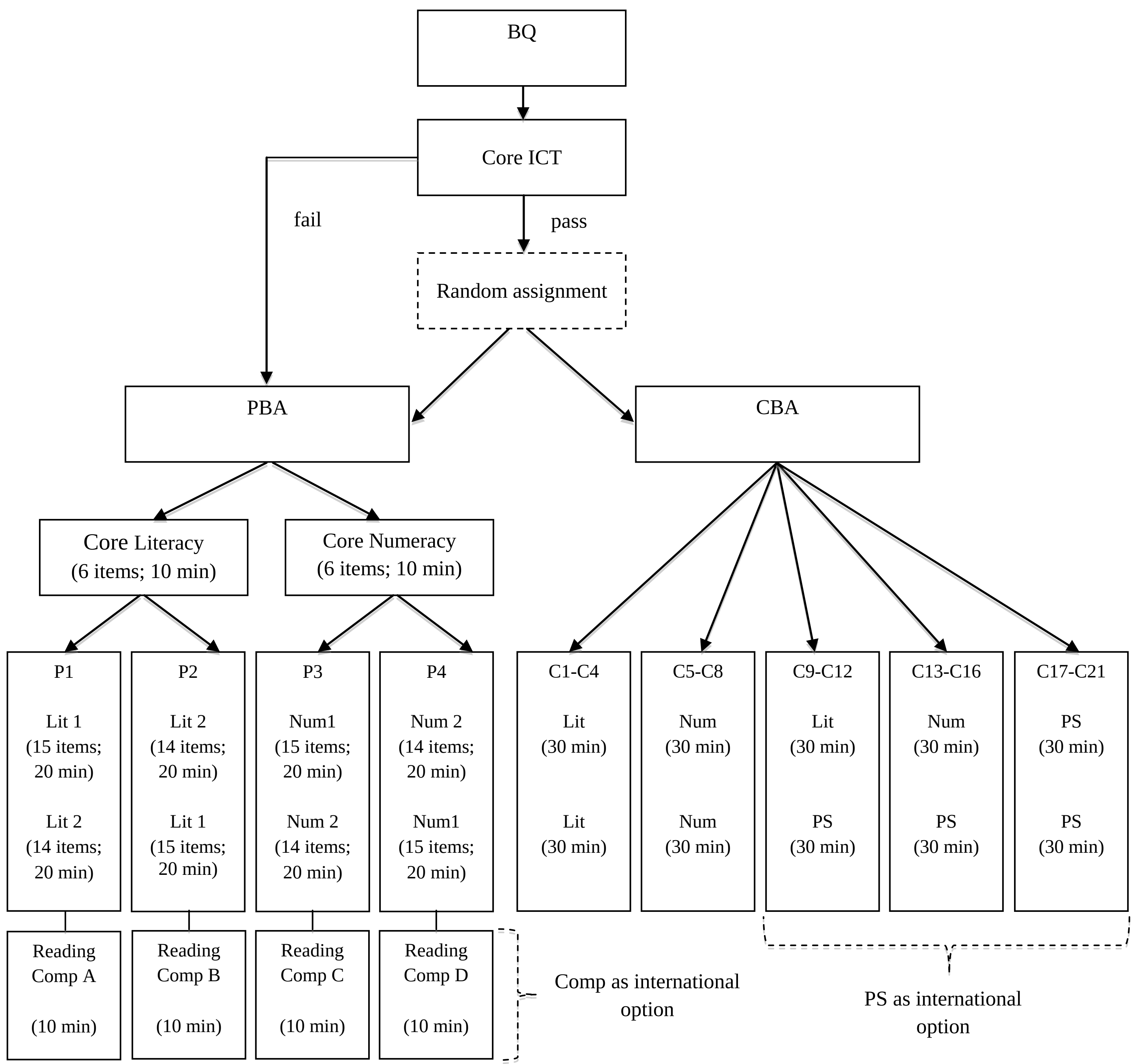
International options in the PIAAC Field Test: The Field Test offered the participating countries the option to assess reading components and problem solving, or not to assess one of them.

The reading components were optional to the participating countries, that is, each country could decide whether to include them in the assessment. Countries choosing the option not to include the reading components measures expected to save about 10 minutes in the overall assessment time and reduced their sample size by a total of 100 adults. The decision not to assess reading components had only minimal impact on the overall Field Test design.

The international option to include reading components but not to assess problem solving had a significant impact on both the sample size needed for the Field Test and the number of computer-based booklets. To compensate for the lack of covariance information between the different domains, the number of respondents per item was increased for the domains of literacy and numeracy, but the overall sample size was smaller (note that the main focus of the Field Test was not on the domain covariance but on the item parameter estimation for each single domain). In this design, assessment time per individual remains at 60 minutes, and each item is answered by 200 adults and based on an estimate of 1,200 respondents per country/per language (i.e., completed cases): 800 who respond to the computer-delivered measures and 400 who respond to the paper-and-pencil items.

In this Field Test design, the direct assessment time was 60 minutes, each item was to be answered by a minimum of 150 adults, and it was based on an estimate of 1,500 respondents per country/per language (i.e., completed cases): 1,100 for the computer-delivered test and 400 for the paper-and-pencil test. Although most countries never reached these numbers, many came close, thus allowing us to carry out the planned analyses.

Figure 19.1: Test design for the PIAAC Field Test 2010



19.2.4 Stage 4: Analyzing Field Test data

Analyses of the Field Test data were carried out to produce overall results as well as results by each participating country. The smallest unit of analysis was language by country data. For the cognitive data, the Field Test analysis included a range of descriptive analyses at both the national and international levels:

- Classical item analysis as well as analyses of collections of items using modern testing methodologies such as IRT
- Analyses of item by survey interaction for common items
- Analyses of item by mode of presentation interaction
- Analyses of item by language within a country
- Selection and rationale supporting the identification of core items, including cut points
- Development of branching rules to be used in the multistage adaptive branching of examinees into different paths of the assessment
- Evaluation of comparability of scoring standard and procedures within and between countries
- Evaluation of anchor booklets (as this was done for the first time in an international large-scale assessment)

The analysis of the Field Test data provided answers to questions related to the finalization of the design of the main assessment as well as the item selection for the main assessment. These questions include the development of the core that, in combination with items from the BQ, guided respondents to the PBA or the CBA and the assembly of booklets and design parameters for the multistage (or adaptive) testing.

The Field Test data were used to examine the comparability of the literacy and numeracy scales for PIAAC against the scales used in IALS and ALL (based on common items across the various surveys). These data were also used to evaluate the stability of the item parameters across the two modes of administration (PBA, CBA). Items that were comparable across the PBA and CBA were used to establish this important link for PIAAC. Field Test data were also used to reveal any item by country interactions and helped quantify these effects, as well as provided information on how they might be reduced (e.g., translation of display issues that can be easily identified and corrected). Results showed several issues associated with clear differences between scoring procedures of PBA and CBA. These findings were incorporated into the improved online scoring during the Main Study and the development of programs to harvest such information from nearly exhaustive log files.

Data on response time were examined as this allowed the Consortium to determine the comparability of time taken on each task across languages/countries, and whether the intended timeframe established in the cognitive labs and previous tryouts hold up as feasible in the Field Test. In addition, the timing of the various blocks and booklets were reviewed and modified.

Item parameters estimated with the Field Test data using IRT analysis were fixed for the adaptive aspect of the Main Study.

19.2.5 Stage 5: Item selection for the Main Study based on the Field Test

The goal of the PIAAC Field Test was to provide new items to cover new domains and extensions of existing frameworks as well as linking items to establish a link among PIAAC, IALS and ALL as well as between PBA and CBA. In order to meet these target goals, it was necessary to develop and assess a larger pool of items for the Field Test compared to the Main Study. The PBA of the Field Test needed a total of 70 items – 35 literacy and 35 numeracy items (while 24+24=48 items were selected for the Main Study). The CBA of the Field Test needed 72 items for each domain (52 items were selected for the Main Study). Of these items, 42 were used to evaluate their utility as linking items for the CBA, while a subset of 25 was used to evaluate their utility for linking the PBA and CBA.

In the Field Test, on average, the respondents from most countries took less time to answer questions than anticipated by nearly 30%. It was decided to lengthen the test by about 10% for Literacy and Numeracy CBA booklets. The reason for not lengthening a full 30% was to reduce the number of respondents going over 60 minutes.

The selection of items for the Main Study was based on three main considerations:

- Measurement construct representations
- Survey design constraints
- Psychometric characteristics of an item as well as a set of items together

The assessment of problem solving in technology-rich environments involved scenarios of varying levels of complexities. Scenarios were designed to take between five and 15 minutes on average to complete. Overall, 150 minutes of testing material was developed for the Field Test (approximately 16 scenarios of varying lengths) with some 75 minutes of problem solving in technology-rich environment tasks selected for inclusion in the Main Study (approximately eight scenarios of varying lengths). The scenarios finally selected for the Main Study were organized into two 25-minute blocks.

With regard to the assessment of reading components, respondents worked through the items more quickly than expected by 2.25 minutes. However, for among least able respondents (below the 17th percentile), the average time was 9.87 minutes. The most able groups of respondents in every country converged to about 3 seconds per item for vocabulary tasks. The proportion correct (P+) differentiated reading components skills of PIAAC respondents rather well for respondents with low skills. For the Main Study, a total of 20 minutes was allotted to measure several of these skills, with final measures assembled from 40 minutes worth of Field Test data.

References

OECD (2012), Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills, OECD Publishing. http://dx.doi.org/10.1787/9789264128859-en

Chapter 20: Creating Simple and Complex Derived Variables and Validation of Background Questionnaire Data

Matthias von Davier, Jonathan Weeks and Henry Chen, ETS;
Jim Allen and Rolf van der Velden, ROA

20.1 Overview

The complex structure of the PIAAC BQ enabled the collection of variables from a diverse population of adults. But not all variables could be reasonably collected for all respondents (e.g., Loeys, Moerkerke, De Smet, & Buysse, 2012). Some were only appropriate for respondents in the workforce, while others were suitable only for those in training. Still another set was used for respondents who belonged to the group of recently unemployed. The need to adapt the BQ in order to provide appropriate sections for a diverse population can be best understood by examining the following examples:

- Current industry and occupation, as well as skill use at work, could be meaningfully asked only of those who were either employed or self-employed at the time of the interview, because respondents who are out of the labor force or never had paid work cannot reasonably be asked whether they use their literacy skills at work.

- For ICT skill use, questions assessing the domain were not presented to those without any previous contact with computers. In contrast, reading, writing and numeracy skills used at home were assessed for all respondents, and the corresponding scales for skills used at work were applied for those respondents who were part of the labor force and the recently (less than 12 months) unemployed.

- Earnings were only asked for those at work. Questions on earnings do not provide meaningful information when respondents are no longer part of the labor force or never had paid work. The same holds true for questions addressing those who were in education or training at the time.

At the same time, a host of other questions in sections addressing general domains are available for practically all respondents who completed the BQ. This is true for skills used at home, education history, questions about health, civil engagement, and approaches to learning, as well as socio-demographic information, among other things. The computer-based routing of respondents to those sections that were appropriate for respondents to answer led to an extremely high item-level response rate overall, as documented in the corresponding section of this chapter.

Clearly, care needs to be taken when analyzing these data. The sections below will provide an overview of some of the key areas for which the Consortium derived variables for use in secondary analyses. The next section presents an overview of those variables that PIAAC shares with previous large-scale assessments of adult populations. The following section discusses the assessment and derivation of earnings variables, and the final section discusses the derivation of variables related to self-reports of literacy skill use, job requirements and learning.

20.2 Overview of the BQ sections

The BQ collected data on a large variety of work-related, education-related and general domains such as socioeconomic variables, health-related questions and attitudinal variables that can be related to the cognitive assessment of literacy skills.

The BQ is too complex to try to reproduce all domains in great detail that were assessed in the instrument. Further information on the development and the content of the BQ is available in Chapter 3 of this report.

A PDF file that provides a linear representation of the international variables collected in the PIAAC BQ can be found at http://www.oecd.org/edu/48442549.pdf.

A framework that outlines the rationale of the selections made in the construction of the different sections of the BQ can be found at http://www.oecd.org/site/piaac/PIAAC(2011_11)MS_BQ_ConceptualFramework_1%20Dec%202011.pdf.

The sections of the BQ broadly covered the following domains relevant for assessing contexts of work, education, skill utilization, and demographics:

A: General information

B: Past education and current education and training

C: Current status and work history

D: Current work (if applicable)

E: Last job (past 12 months if no current job)

F: Skills used at work (JRA)

G: Literacy, numeracy, ICT at work

H: Literacy, numeracy, ICT at home

I: About yourself

J: Background

As stated above, the path through the BQ was an adaptive one, as different sections were appropriate for respondents who were employed, unemployed, out of the labor force, or still in school or training. Altogether, there were over 400 questions (without national adaptation), so it becomes virtually impossible to report in detail on each of the questions. Instead, we provide Table 20.1, which shows the rate of response by country for those adaptively routed question paths presented to respondents. That is, only respondents that received questions are counted in terms of response or nonresponse.

Table 20.1: Response rate in detail

Country	**BQ Rate**	**MIN**	**MAX**	**MEAN**	**Median**	**MIN Item**	**No Response**	**Below 50%**	**50% to 90%**
Australia	88.2%	88.0%	99.5%	98.9%	99.4%	J_Q07b	37	0	2
Austria	50.3%	6.1%	100.0%	98.8%	99.8%	B_D01d	0	3	0
Canada	80.6%	9.9%	99.5%	97.9%	98.7%	B_D01d	0	2	0
Cyprus[1]	86.6%	7.8%	100.0%	98.7%	99.5%	B_D01d	0	2	0
Czech	40.5%	9.7%	100.0%	99.1%	99.9%	B_D01d	0	2	0
Denmark	46.3%	11.8%	99.6%	98.4%	99.1%	B_D01d	1	2	0
England/ N. Ireland	68.5%	15.0%	100.0%	99.1%	99.9%	B_D03d	0	2	2
Estonia	58.4%	7.5%	100.0%	98.9%	99.7%	B_D01d	0	2	0
Finland	70.7%	8.8%	99.9%	94.7%	95.3%	B_D01d	0	2	1
Flanders (Belgium)	54.2%	6.8%	100.0%	99.0%	99.9%	B_D01d	0	2	1
France	72.5%	8.3%	100.0%	90.2%	90.9%	B_D01d	0	2	18
Germany	52.5%	9.8%	100.0%	98.9%	99.9%	B_D01d	0	2	1
Ireland	92.7%	9.2%	99.9%	99.0%	99.8%	B_D01d	0	2	1
Italy	62.7%	4.5%	100.0%	98.4%	99.2%	B_D01d	0	2	0
Japan	47.0%	5.2%	100.0%	99.0%	99.9%	B_D01d	0	2	1
Korea	78.4%	9.9%	99.1%	97.7%	98.4%	B_D01d	0	2	0
Netherlands	50.3%	7.9%	100.0%	99.1%	99.8%	B_D01d	0	2	0
Norway	58.9%	12.6%	99.9%	98.9%	99.8%	B_D01d	0	2	1
Poland	51.0%	20.5%	98.0%	97.0%	97.7%	B_D01d	0	2	0
Russian Fed.[2]	99.5%	15.0%	100.0%	98.9%	99.8%	B_D01d	0	2	2
Slovak	61.6%	5.2%	100.0%	99.0%	99.8%	B_D01d	0	2	0
Spain	41.5%	11.5%	100.0%	99.1%	99.8%	B_D01d	0	2	0
Sweden	44.8%	10.5%	100.0%	99.0%	99.8%	B_D01d	0	2	0
United States	80.6%	10.0%	99.9%	98.7%	99.6%	B_D01d	0	2	2

[1] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[2] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 20.1 shows the mean and median percentage response rate, as well as the minimum and maximum, along with the item for which the minimum was observed. The last three columns provide an overview of the number of items without any responses, and those with responses below 50% and between 50% and 90%. It can be seen that due to confidentiality deletions, a few countries exhibit nonzero counts for items without any responses.

20.3 Overview of BQ trend variable domains

One of the major tasks of international assessments is providing trend information. For that reason, the PIAAC Consortium tried to collect and derive variables that can be viewed as comparable over three adult assessments: IALS, ALL and PIAAC. In order to achieve this, the Consortium developed a number of derived variables (DVs) based on the raw BQ variables collected during the computer-based background interview.

Table 20.2 shows an overview of these indices. It can be seen that several variables have a direct correspondence among the three assessments, while there are also variables that required derivation for one or more of the three assessments in order to arrive at a comparable definition across all three assessments. In order to do so, some variables had to be coarsened for the purpose of defining a variable that allows quantitative comparisons across assessments based on groupings using trend variables.

Table 20.2: Domains with available trend variables*

	Domain
1	Date of birth
2	Gender
3	Respondent’s origin
4	Educational background - formal education
5	Language background
6	Respondent’s mother’s background
7	Respondent’s father’s background
8	Respondent’s employment status
9	Work history - past 12 months
10	Job information - current job or last (past 12 months) job held
11	Education or training which the respondent has taken in the past 12 months
12	Education or training wanted but not taken in the past 12 months
13	Reading and writing in respondents’ daily life
14	Civic participation - volunteer work
15	Health
16	Use of information technologies - computer use
17	Respondents’ children’s education

* Detailed information about matching variables in the BQ across IALS, ALL and PIAAC instruments is given in Appendix 4

While not all domains include trend variables, many of the central reporting variables were able to be matched. If no direct match could be achieved, questions that largely agree were identified. Appendix 4 provides the details on questions that are matched between IALS, ALL and PIAAC.

20.4 Development of derived earnings variables

20.4.1 Introduction

The BQ deployed two key innovations designed to make it easier for respondents to report their earnings, and thereby to improve the quality of the available earnings data and reduce item nonresponse. The first was for respondents to choose among reporting their earnings per hour, day, week, two weeks, month or year, or by piece rate. By removing the necessity for respondents to convert from their own preferred payment period to a predetermined standard, the aim was to improve the data quality and remove potential barriers to response. Furthermore, this approach automatically takes into account country differences in the payment period that are typically applied in most cases.

The second key innovation was an additional option for those who were still unwilling or unable to report their earnings as a precise amount. In this case, respondents were invited to report their earnings in broad categories. Again, in this case the categories were expressed per hour, per day, per week, per two weeks, or per month or per year according to the respondent's preference. This option was expected to be attractive for respondents who had only a rough idea of how much they earn per period and for those reluctant to reveal their precise earnings due to concerns such as privacy.

In addition to these key innovations, earnings were asked separately for wage and salary earners and for the self-employed, and there was a separate question for wage and salary earners in which they could report annual bonuses they may have received. Earnings of self-employed were asked per year, unless respondents had been in their current business for less than a year, in which case they were asked per month. For earnings of both wage and salary earners and the self-employed, as well as for annual bonuses, the option to report in broad categories was offered for those who were unwilling or unable to report directly.

Although the design of the set of questions was expected to yield significant advantages in terms of interview flow, item response and data quality, these advantages come at a price—there is no direct measure of earnings that ensues directly from the data. It was necessary to devise a fairly elaborate set of conversion rules to go from earnings as reported to the derived earnings variables used in the data. The first step is a fairly straightforward conversion of directly reported earnings from the earnings period option chosen by the respondent into every available alternative (e.g., from hourly to monthly, from yearly to daily, etc.). The second, and by far most complex, step comprised the conversion of earnings reported in broad categories into an equivalent direct amount. A third step comprised the construction of a set of standard variables that formed the basis for the earnings derived variables (DVs) to be included in the public data file. A fourth step involved a purchasing power parity (PPP) correction so that all earnings variables were expressed in terms of real disposable earnings in a fixed currency (in this case given in US dollars). Finally all earnings indicators were converted into deciles.

20.4.2 Conversion of directly reported earnings into all possible reporting periods

As stated above, this step was quite straightforward and involved using a set of fixed conversion rules from each reporting period into every other reporting period; earnings reported as a piece rate were first converted into an hourly rate based on an additional question regarding the usual number of hours per piece as estimated by the respondent. This conversion makes use of the number of hours worked per week, using rules on the ratio between the different reporting periods. Most of these variables are not intended for inclusion in the final data, which only include earnings expressed in hourly or monthly amounts. The reason for creating all of these variables is that they are needed as input for the following step, the conversion of earnings reported in broad categories into an equivalent direct amount.

20.4.3 Converting broad categories into equivalent direct amounts

As stated above, any respondents who were unable or unwilling to report their earnings precisely were given the option of reporting in broad categories. These categories were provided by each participating country on the basis of their national earnings distribution. For regular earnings of wage and salary earners, six broad bands were used, with the bands divided roughly along the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles of the national distribution, provided separately per hour, day, week, two weeks, month or year. For self-employed, the same bands were applied, but only per year or month, depending on whether the respondent had been in the current business at the time of the survey for at least a year or less than a year. For annual bonuses, three broad bands were used, with bands divided at roughly 5% and 10% of the median of national annual gross earnings.

Convenient as this option may have been for some respondents, it does not yield a unique earnings amount that can be directly compared with the direct earnings reported by the majority of respondents. Several alternative approaches were considered for dealing with this problem:

- Replacing the bands by a fixed amount, for example, the midpoint of the band or some other value considered to be the most likely value. This option was rejected for a number of reasons. The most important reason is that this would give rise to unwanted "lumpiness" in the data, which is not only a problem in its own right but leads to unavoidable and unsolvable problems when converting final earnings into deciles in a later step. Conversion into six discrete amounts inevitably means that all earnings reported in broad categories would be included in just six of the deciles. A further complication of this approach was caused by the fact that the broad bands were not usually strictly comparable across reporting periods. This was because countries usually rounded the dividing points into round amounts, for example, 6 Euros rather than, say, 5.78 Euros, which might be a strict conversion from the equivalent dividing point in terms of monthly earnings.

- Converting direct earnings into the six broad bands. This option was rejected for reasons similar to the previous option. In addition to the above-mentioned discrepancies between the different reporting periods within each country, there was the additional problem that there are non-negligible differences between the manner in which the bands were defined per country, which would negatively affect comparability. Finally, it was observed that the bands used in the BQ were never intended to be used in this way, and in fact represent a highly unusual way in which to express earnings.

- Leaving the data unconverted, allowing users of the data to make their own conversions as they see fit. This option was not seriously considered, because it would essentially render the earnings data as included in the public data file for this group of respondents unusable.

Taking into account the serious limitations of the alternatives considered above, it was decided that a precise earnings amount would be imputed for every respondent who reported in broad categories. The imputation method comprised matching each of these respondents with a respondent who reported earnings directly, meaning the person was considered "most likely" to resemble him or her in terms of earnings, and assigning the precise amount reported by that respondent. The basis for this matching was predicted earnings on the basis of a regression model using key indicators such as highest education, skill level, age, gender and so on as predictors.

In somewhat more detail, the imputation process followed the following steps:

1. Precise earnings of wage and salary earners were converted into the same broad ranges as used in the BQ.
2. Earnings regressions were run on directly reported earnings, separately for hourly, daily, weekly, biweekly, monthly and yearly earnings, in each case also separately for low, medium and high earnings (earnings bands 1-2, 3-4 and 5-6 respectively).
3. Predicted earnings were saved in each case, both for those who reported earnings directly and those who reported in broad categories.
4. Cases that reported in broad categories were matched to their "nearest neighbor" in terms of predicted earnings among those who reported directly. This matching was conducted separately for each of the broad earnings ranges, thus ensuring that each case would always be matched with a "mate" who fell into the same broad category.
5. Based on this matching, each broad category case was assigned the actual directly reported hourly and monthly earnings value of its "mate." Note that this assignment always takes place based on the matching based on the reporting category actually used by respondents. For example, those who reported earnings based on an hourly rate were always matched on the basis of predicted hourly earnings. In this way, we ensured that the matching was as precise as possible, and removed any possible bias that might occur because the dividing amounts for the different reporting periods were not always strictly equivalent.
6. An equivalent process was used to derive imputed values for additional payments.
7. The imputed hourly and monthly earnings, as well as the imputed additional payments, were combined with directly reported earnings to form single hourly and monthly earnings variables.
8. A flag variable was created to indicate whether earnings were imputed or directly reported.

It should be noted that it did not prove possible to derive imputed earnings for the self-employed using such a methodology. The primary reason is the unusual earnings distribution for self-employed, and in particular the fact that a significant proportion of the self-employed had zero or negative earnings (in both cases reported as zero in the data). We were not successful in developing sufficiently reliable and robust regression models that were able to account for the unusual composition of this category of the self-employed, in terms of education, skills and other factors.

20.4.4 Construction of a set of standard variables

Starting with the above mentioned variables for wage and salary earners, combining actual and imputed earnings (hourly and monthly earnings, additional payments) and the direct monthly earnings measure for the self-employed, we then constructed a set of standard variables that formed the basis for the earnings DVs to be included in the public data file. The first two of these were hourly and monthly earnings of wage and salary earners, excluding bonuses. By adding additional payments to these (of course, with the necessary conversion to the payment period concerned), we then constructed two variables comprising hourly and monthly earnings of wage and salary earners including bonuses. By combining monthly earnings of wage and salary earners including bonuses with monthly earnings of the self-employed, we obtained an overall measure of total monthly earnings of wage and salary earners and self-employed.

20.4.5 Purchasing power parity (PPP) conversion

The next step involved a PPP correction, so that all earnings variables were expressed in terms of real disposable earnings in a fixed currency (in this case, US dollars). This is simply a multiplication by a constant value per country, based on data on purchasing power parity per country supplied by the OECD.

20.4.6 Conversion into deciles

Finally all earnings indicators were converted into deciles. This involved dividing the data for each earnings variable into 10 equally sized groups per country strictly based on the position in the distribution of earnings according to that variable. Where there were multiple cases at the cutoff points, these respondents were assigned to the higher and lower earnings group in the numbers required to produce groups of equal size, with individuals being randomly sorted into the higher and lower groups.

20.5 Derivation of variables related to self-reports of literacy skill use, job requirements and learning

20.5.1 Overview

In PIAAC, the skills of a population are not only measured directly through the cognitive instruments but also indirectly through the BQ by asking respondents to report on their use of skills both inside and outside of work. The frequency and type of activities associated with reading, writing, numeracy and information technology were targeted in the BQ using multiple items that were similarly worded to apply to activities both in and out of work. In addition, other areas, particularly those involving intrapersonal, interpersonal and other generic “soft” skills, not included in the direct assessment, were also addressed through a set of self-reported questions.

This set of questions makes up a module within the questionnaire that has been specifically developed for the PIAAC project: the JRA module.

Altogether, we constructed 13 scales based on a cross-country analysis of comparability, reliability, and convergent as well as discriminant validity. These scales were constructed using IRT, more specifically, the generalized partial credit model (GPCM), and person-specific levels of skill use were estimated using weighted likelihood estimation (WLE). Scale values were derived for all respondents who reported at least some limited activities in each of these domains. Those who reported no skill use in each of the 13 areas were not represented on any of these 13 scales; nevertheless, they provide important information with respect to the percentage of people in each participating country who do not use particular type of skills either in or outside of work. The data also allow users to examine the characteristics of these individuals.

20.5.2 Models and methods

The PIAAC BQ contains scales, that is, collections of questions around a topic, with regards to the domains of skill use, activities at work, and approaches to learning. These scales are mainly found in sections F, G, H and I of the BQ. The skill use scales are arranged around domains that relate to the literacy domains assessed in the cognitive part (the test) of the PIAAC. More specifically, questions around activities involving reading, writing, numeracy, and the use of technology were administered and respondents were asked to rate how often they perform these activities either at work (section G) or outside of work (section H).

Based on the arrangement of these questions within sections, and based on the topics covered therein, there are a number of groupings for these BQ items that a reviewer may come up with. Upon OECD's request, the PIAAC Consortium tested a set of 30 potential scales. The question is whether a given set of items can provide reliable, nonredundant measures of skill use (and other behavioral indicator) to justify the reporting of these results as a derived indicator. Three principal criteria were used to determine if a specific scale should be retained: average internal consistency reliabilities (as measured by Cronbach's alpha) across countries greater than or equal to 0.6, mean subscale (total score) correlations across countries less than 0.7, and ignorable country misfit as characterized by weighted root mean squared differences between empirical and expected response probabilities across countries. In all cases, the number of items associated with each potential scale is quite small (two to eight items); hence, the ability estimator must also be considered so as to minimize bias. The most common estimators of latent ability are maximum likelihood (ML) and expected a posteriori (EAP). The former does not incorporate any bias correction whereas the latter is a Bayesian approach that shrinks estimates toward the mean as a function of score reliability. In order to minimize bias without reducing the variability of the scores considerably, a weighted maximum likelihood estimation (WLE) approach was used.

20.5.2.1 Item parameter estimation

The skill use items as well as the items used in the approaches to learning and the job requirement analyses are measured using a five-point Likert scale. The items for each potential scale were fitted using the generalized partial credit model (GPCM; Muraki, 1992).[3]

[3] For the potential scales with only two items, a partial credit model (Masters, 1982) was used where the item slopes were constrained to be the same for both items.

$$P_{ij}(\theta) = \frac{\exp\left[\sum_{c=0}^{j} Da_i(\theta - b_{ic})\right]}{\sum_{c=0}^{J} \exp\left[\sum_{k=0}^{c} Da_i(\theta - b_{ik})\right]} \tag{1}$$

where the probability of responding in a given category, k, is modeled as a function of examinee ability, θ, and estimated item parameters. For the GPCM, a_i is the discrimination (slope) for item i, b_{ik} is a step parameter. D is a scaling constant equal to 1.7, not an estimated parameter, which is included in the estimation for reasons that relate the logistic models to probit models (e.g. Cramer, 2004). The item parameters were estimated using the PARSCALE software, implementing a multiple-group concurrent calibration with countries serving as the different groups, and countries equally weighted by means of standardizing the sampling weights to a constant sum per country. The estimation utilizes marginal maximum likelihood without any priors specified for the item parameters and is run in two phases. In the first phase the mean and standard deviation of the ability scale are fixed to 0 and 1 respectively, across countries, and the ability distribution is constrained to be normal. In the second phase the item parameter estimates from the first phase are used as starting values and the estimation proceeds after relaxing the normality constraint.

Scale Exclusion Criteria

With the item parameters estimated, the BQ items were checked for misfit and each of the potential scales was considered for elimination. Three primary exclusion criteria were used to identify items/scales that were problematic and/or provided redundant information:

Criterion 1: Scale Reliability – When reporting subscale results, it is important that the scores have sufficient reliability to allow for defensible inferences to be made on the basis of the scores. For cognitive measures, reliabilities of 0.70 or higher are generally preferred. If this criterion were used, nearly two-thirds of the potential scales would be flagged for possible exclusion. As such, a slightly relaxed criterion was used. In order to be considered for exclusion, the mean reliability across countries had to be less than 0.6, as characterized by Cronbach's alpha.

Criterion 2: Scale Correlations – In addition to being reliable, subscores should provide unique information about the measured background characteristics. Scales that provide redundant information may be of little utility; hence, the correlation between scales was considered. Potential scales with a mean correlation across countries greater than or equal to 0.7 were flagged for possible exclusion.

Criterion 3: Between Country Differences – When item parameters are estimated for measures administered across countries, there is a strong potential for item-by-country interactions which may lead to item misfit. Stated differently, the empirical response curves across countries may differ appreciably from the expected curves based on international item parameters. These differences may occur for individual items or all/most items in a subscale. To summarize these differences, a weighted root mean squared difference (WRMSD)

$$WRMSD_i = \sqrt{\sum_{C}\sum_{X} \frac{\omega_c(X)[p_{ijc}(X) - P_{ij}(X)]^2}{J}} \tag{2}$$

can be computed for each BQ item, i, where $P_{ij}(X)$ is the expected probability of responding in category j for a given ability, X, $p_{ijc}(X)$ is the proportion of examinees in country C responding in category j, and $\omega_c(X)$ is a set of weights corresponding to the expected proportion of examinees in country C at ability X for the given subscale. Items with a WRMSD greater than 0.25 logits were flagged for possible exclusion. Additionally, scales where more than half of the items had WRMSDs greater than 0.25 were flagged for possible exclusion.

20.5.2.2 Skill use level estimation

Once the final set of subscales was identified, examinee skill use levels for each scale were estimated. All of the potential scales have very few items; hence, there is an increased potential for bias in estimates of ability. The most common estimators of latent ability are ML and EAP. As mentioned earlier, the former does not incorporate any bias correction whereas the latter is a Bayesian approach that shrinks estimates toward the mean as a function of score reliability. As an alternative to EAPs, Warm (1989) proposed a weighted likelihood estimator for dichotomously scored responses that essentially serves as a bias-corrected ML estimator. Penfield and Bergeron (2005) extended this methodology to GPCM items.

The maximum likelihood estimate of θ for a given individual is equal to the value of θ that maximizes the log likelihood, L, of the associated response pattern given a fixed set of item parameters. This estimate is obtained iteratively through the use of the Newton-Raphson algorithm where the estimate at iteration t is equal to

$$\hat{\theta}_t = \hat{\theta}_{t-1} - \frac{L'}{L''} \tag{3}$$

In Equation 3, L' and L'' are given by

$$L' = \sum_{i=1}^{N} \sum_{j=0}^{J} u_{ij} D a_i (j - \lambda_1), \tag{4}$$

$$L'' = -\sum_{i=1}^{N} D^2 a_i^2 (\lambda_2 - \lambda_1^2), \tag{5}$$

where $\lambda_k = \sum_{j=0}^{J} j^k P_{ij}$ and P_{ij} is the expected probability from (1). Under this formulation $\lambda_1 = \sum_{j=0}^{J} j P_{ij}$ and $\lambda_2 = \sum_{j=0}^{J} j^2 P_{ij}$. The standard error for $\hat{\theta}$ is equal to

$$SE(\hat{\theta}) = \frac{1}{\sqrt{I}} \tag{6}$$

where I is the information of the test at θ, and is computed as

$$I = \sum_{i=1}^{N} D^2 a_i^2 (\lambda_2 - \lambda_1^2) \tag{7}$$

Extending this approach, the weighted likelihood estimator of θ at iteration t is equal to

$$\hat{\theta}_t = \hat{\theta}_{t-1} - \frac{W'}{W''} = \hat{\theta}_{t-1} - \frac{L' + B'}{L'' + B''} \tag{8}$$

where W is the weighted log likelihood (i.e., the bias corrected log-likelihood) and B' and B'' are given by

$$B' = \frac{\sum_{i=1}^{N} D a_i^3 (\lambda_3 - 3\lambda_1\lambda_2 + 2\lambda_1^3)}{2\sum_{i=1}^{N} a_i^2 (\lambda_2 - \lambda_1^2)} \tag{9}$$

$$B'' = \frac{AB - 2C^2}{B^2} \tag{10}$$

where

$$A = \sum_{i=1}^{N} D^2 a_i^4 (\lambda_4 - 4\lambda_1\lambda_3 - 3\lambda_2^2 + 12\lambda_1^2\lambda_2 - 6\lambda_1^4) \tag{11}$$

$$B = 2\sum_{i=1}^{N} a_i^2 (\lambda_2 - \lambda_1^2) \tag{12}$$

$$C = \sum_{i=1}^{N} D a_i^3 (\lambda_3 - 3\lambda_1\lambda_2 + 2\lambda_1^3) \tag{13}$$

Because B is proportional to the likelihood, it cannot be estimated directly (Warm, 1989). As such, the Newton-Raphson method is required. The standard error is the same as that obtained for the ML estimate.

20.5.3 Potential scales

By clustering related BQ items, 30 potential scales were identified by OECD analysts and the Consortium was asked to evaluate these scales. This list of scales included 18 non-nested scales and 12 nested scales (comprising subsets of items from four non-nested scales). In the list below, the values in the parentheses indicate the number of items associated with each scale.

Non-nested scales:

- Cooperation (2)
- ICT at home (7)
- ICT at work (7)
- Influence (7)
- Learning at work (3)
- Numeracy at home (6)

- Numeracy at work (6)
- Physical (2)
- Planning (3)
- Problem solving (2)
- Reading at home (8)
- Reading at work (8)
- Readiness to learn (6)
- Self-organization (2)
- Task discretion (4)
- Trust (2)
- Writing at home (4)
- Writing at work (4)

Nested Scales:

- Numeracy at home: Basic (3), Advanced (3)
- Numeracy at work: Basic (3), Advanced (3)
- Reading at home: Basic (3), Advanced (5); Documents (4), Prose (4)
- Reading at work: Basic (3), Advanced (5); Documents (4), Prose (4)

20.5.4 Results

In an effort to provide only scale-based derived variables that meet a sufficient level of psychometric quality, all proposed scales were analyzed first for each of the participating countries separately, and then jointly for consistency across countries. While scales with two items are viewed with well-grounded concern (Eisinga, Te Grotenhuis, & Pelzer, 2013), they were included in this first round of analyses in order to ensure that all of the proposed scales would be checked as requested.

20.5.4.1 Scale reliabilities

Table 20.3 presents the mean, standard deviation, minimum, and maximum of the country-level reliabilities for each potential scale. The mean reliabilities ranged from 0.50 to 0.84 for the non-nested scales and 0.50 to 0.78 for the nested scales. Using the criterion of alpha values less than 0.6, three non-nested scales and four nested scales were flagged for possible exclusion. Three of these scales had mean alpha values substantively below 0.6: physical ($r = 0.49$), reading at home: basic ($r = 0.50$), and writing at home ($r = 0.51$). The other four scales had mean reliabilities at or slightly below 0.6: cooperation ($r = 0.59$), reading documents at home ($r = 0.60$), reading prose at home ($r = 0.58$), and reading documents at work ($r = 0.60$).

Table 20.3: Reliability summary statistics for potential subscales

		Mean	SD	Min	Max
Non-Nested Scales	Cooperation	**0.59**	0.07	0.48	0.70
	ICT at home	0.69	0.03	0.64	0.76
	ICT at work	0.77	0.02	0.73	0.81
	Influence	0.79	0.02	0.74	0.82
	Learning at work	0.69	0.05	0.59	0.80
	Numeracy at home	0.77	0.03	0.72	0.82
	Numeracy at work	0.81	0.02	0.77	0.84
	Physical	**0.49**	0.22	-0.26	0.71
	Planning	0.71	0.04	0.62	0.77
	Problem solving	0.68	0.04	0.57	0.74
	Reading at home	0.73	0.04	0.66	0.80
	Reading at work	0.81	0.03	0.75	0.85
	Readiness to learn	0.84	0.03	0.80	0.91
	Self organization	0.79	0.08	0.53	0.88
	Task discretion	0.80	0.04	0.73	0.92
	Trust	0.66	0.07	0.46	0.80
	Writing at home	**0.50**	0.05	0.37	0.63
	Writing at work	0.62	0.06	0.51	0.77
Nested Scales	Numeracy at home: Basic	0.68	0.04	0.61	0.73
	Numeracy at home: Adv	0.72	0.08	0.56	0.81
	Numeracy at work Basic	0.79	0.04	0.67	0.84
	Numeracy at work Adv	0.68	0.07	0.53	0.76
	Reading at home: Basic	**0.50**	0.09	0.36	0.67
	Reading at home: Adv	0.62	0.04	0.56	0.69
	Reading at home: Docs	**0.60**	0.04	0.53	0.66
	Reading at home: Prose	**0.58**	0.07	0.46	0.71
	Reading at work: Basic	0.68	0.05	0.57	0.77
	Reading at work: Adv	0.67	0.04	0.59	0.76
	Reading at work: Docs	**0.60**	0.05	0.54	0.69
	Reading at work: Prose	0.78	0.03	0.70	0.83

20.5.4.2 Scale correlations

Table 20.4 presents the mean raw-score correlation between the potential subscales. Using the criterion of correlations greater than or equal to 0.7, there are three sets of scales that appear to provide redundant information. These sets correspond primarily to the non-nested scales. The only exception is for the subscales for *self-organization* and *planning,* which were strongly correlated across all countries (mean r = 0.91).

The scales for *reading skills at home* for both document and prose type texts, and the scales for basic and advanced literacy skills at home (i.e., the nested scales for reading at home) generally had high moderate to high correlations across all countries (range of mean correlations: 0.79 – 0.93). Similarly, the scales for *reading skills at work* for both document and prose type texts, and the scales for basic and advanced literacy skills at work (i.e., the nested scales for reading at work) generally had moderate to high correlations across all countries (range of mean correlations: 0.70 – 0.94).

The subscale for *numeracy at home* was strongly correlated with both basic and advanced numeracy at home across all countries (range of mean correlations: 0.84 – 0.91), yet basic and advanced numeracy at home were only moderately correlated (r = 0.53). The subscale for *numeracy at work* was strongly correlated with both basic and advanced numeracy at work across all countries (range of mean correlations: 0.83 – 0.92), but basic and advanced numeracy at work were only moderately correlated (r = 0.56).

Table 20.4: Subscale correlations, averaged across countries

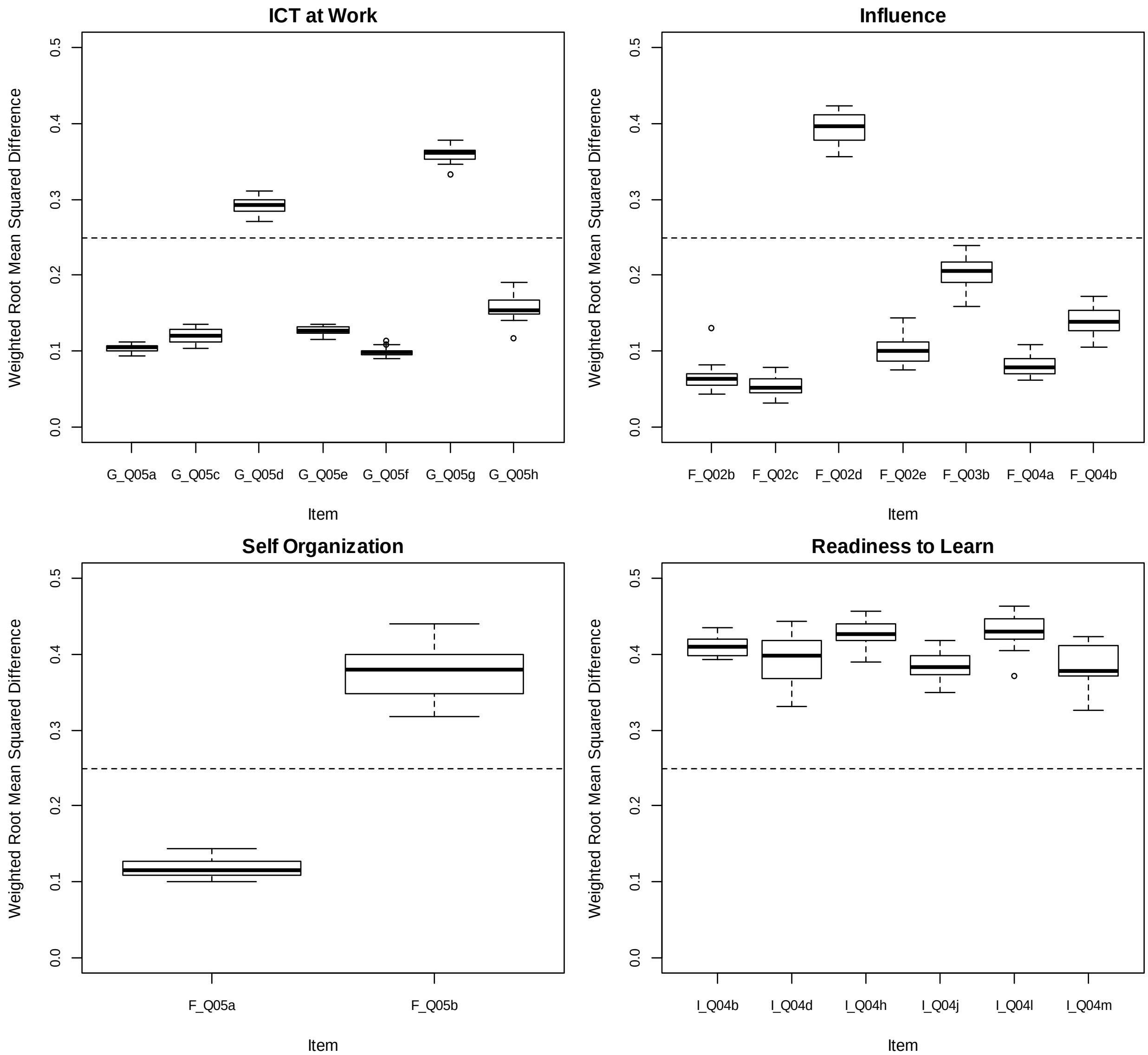
	COOPERATION	ICTHOME	ICTWORK	INFLUENCE	LERNATWORK	NUMHOME	NUMHOMEADV	NUMHOMEBAS	NUMWORK	NUMWORKADV	NUMWORKBAS	PHYSICAL	PLANNING	PROBWORK	READHOME	READHOMEADV	READHOMEBAS	READHOMEDOC	READHOMEPRO	READWORK	READWORKADV	READWORKBAS	READWORKDOC	READWORKPRO	READYTOLERN	SELFORGANISE	TASKDISC	TRUST	WRITHOME
ICTHOME	0.05																												
ICTWORK	0.02	0.37																											
INFLUENCE	0.25	0.18	0.34																										
LERNATWORK	0.27	0.17	0.21	0.36																									
NUMHOME	0.03	0.45	0.19	0.18	0.15																								
NUMHOMEADV	0.02	0.41	0.17	0.11	0.10	0.84																							
NUMHOMEBAS	0.04	0.38	0.17	0.18	0.15	0.91	0.53																						
NUMWORK	0.08	0.24	0.54	0.48	0.26	0.35	0.28	0.32																					
NUMWORKADV	0.09	0.26	0.50	0.37	0.22	0.36	0.36	0.28	0.83																				
NUMWORKBAS	0.40	0.18	0.44	0.47	0.24	0.27	0.17	0.28	0.92	0.56																			
PHYSICAL	0.10	-0.12	-0.32	-0.10	-0.02	-0.04	-0.06	-0.02	-0.21	-0.21	-0.17																		
PLANNING	0.14	0.11	0.36	0.62	0.25	0.12	0.06	0.13	0.39	0.33	0.36	-0.09																	
PROBWORK	0.21	0.18	0.35	0.47	0.37	0.17	0.11	0.17	0.39	0.35	0.35	-0.11	0.41																
READHOME	0.07	0.51	0.28	0.31	0.23	0.51	0.39	0.49	0.32	0.31	0.26	-0.11	0.23	0.28															
READHOMEADV	0.05	0.44	0.26	0.30	0.21	0.49	0.39	0.47	0.30	0.31	0.24	-0.10	0.23	0.26	0.93														
READHOMEBAS	0.08	0.45	0.22	0.24	0.20	0.40	0.28	0.40	0.25	0.23	0.21	-0.08	0.18	0.23	0.83	0.57													
READHOMEDOC	0.06	0.42	0.21	0.25	0.20	0.53	0.40	0.51	0.30	0.29	0.24	-0.04	0.19	0.24	0.86	0.83	0.67												
READHOMEPRO	0.06	0.45	0.26	0.29	0.21	0.38	0.29	0.37	0.26	0.26	0.22	-0.15	0.22	0.25	0.89	0.80	0.79	0.54											
READWORK	0.14	0.26	0.61	0.58	0.40	0.24	0.16	0.24	0.61	0.54	0.54	-0.23	0.50	0.51	0.46	0.43	0.37	0.38	0.41										
READWORKADV	0.10	0.25	0.55	0.55	0.37	0.25	0.18	0.25	0.61	0.55	0.53	-0.19	0.47	0.47	0.44	0.44	0.33	0.39	0.39	0.94									
READWORKBAS	0.16	0.21	0.55	0.52	0.37	0.18	0.10	0.19	0.50	0.44	0.45	-0.23	0.45	0.47	0.39	0.35	0.35	0.31	0.37	0.90	0.70								
READWORKDOC	0.15	0.22	0.48	0.50	0.37	0.24	0.16	0.25	0.61	0.52	0.55	-0.13	0.42	0.47	0.39	0.37	0.31	0.38	0.31	0.89	0.88	0.74							
READWORKPRO	0.09	0.24	0.57	0.54	0.35	0.19	0.12	0.19	0.50	0.46	0.43	-0.27	0.47	0.45	0.43	0.41	0.35	0.32	0.43	0.91	0.82	0.87	0.62						
READYTOLERN	0.09	0.34	0.25	0.31	0.27	0.31	0.23	0.30	0.27	0.26	0.23	-0.08	0.25	0.28	0.44	0.42	0.35	0.36	0.40	0.34	0.33	0.29	0.28	0.32					
SELFORGANISE	0.05	0.09	0.34	0.44	0.21	0.10	0.04	0.12	0.35	0.28	0.32	-0.10	0.91	0.36	0.20	0.19	0.16	0.16	0.20	0.43	0.40	0.40	0.37	0.41	0.22				
TASKDISC	-0.09	0.12	0.32	0.30	0.13	0.11	0.09	0.10	0.32	0.23	0.32	-0.16	0.43	0.21	0.16	0.16	0.12	0.13	0.15	0.32	0.31	0.26	0.25	0.31	0.23	0.45			
TRUST	0.02	0.08	0.11	0.10	0.05	0.06	0.06	0.04	0.08	0.10	0.06	-0.14	0.10	0.09	0.15	0.15	0.10	0.09	0.16	0.14	0.13	0.13	0.08	0.17	0.07	0.09	0.08		
WRITHOME	0.05	0.57	0.26	0.25	0.19	0.48	0.41	0.43	0.25	0.25	0.20	-0.11	0.16	0.20	0.60	0.53	0.54	0.52	0.54	0.31	0.30	0.27	0.26	0.30	0.36	0.13	0.12	0.11	
WRITWORK	0.16	0.24	0.54	0.49	0.30	0.18	0.12	0.19	0.51	0.47	0.43	-0.24	0.43	0.45	0.36	0.33	0.30	0.28	0.34	0.69	0.61	0.67	0.60	0.64	0.28	0.37	0.22	0.13	0.33

20.5.5 Between-country differences

Figure 20.1 presents box-and-whiskers plots of the WRMSDs for each item for four of the potential subscales. These types of plots were used to visually identify potentially problematic items/subscales. Out of the full set of BQ items, there were 20 items with WRMSD values greater than 0.25 logits. These results point to items that function differentially across countries. In most instances, these were single items on a given scale; however, there were five cases where all, or the majority, of items associated with a given scale had WRMSD values greater than or equal to the criterion. These scales (with mean WRMSDs in the parentheses) include:

cooperation (0.29), *physical* (0.31), *problem solving* (0.25), *readiness to learn* (0.41) and *trust* (0.44). Most of these are two-item scales.

Figure 20.1: Subscale weighted root mean squared differences



20.5.6 Subscale retention determinations

Based on the results of these analyses, a decision was made to exclude all two-item and nested scales. A total of 13 subscales were retained. Each of the two-item scales was flagged for exclusion based on one or more criteria. *Cooperation* and *physical* were flagged as problematic both for low reliability and between country differences. *Problem solving* and *trust* were flagged as problematic due to between-country differences, and *self-organization* was strongly correlated with the three-item scale *planning*. All four of the nested scales were highly correlated with the corresponding non-nested scales, indicating that the nested subscales provided redundant information relative to the associated non-nested scales. The subscales for reading documents/prose at home and work also had low reliabilities. In addition to the exclusion of these scales, two items were eliminated due to large between-country differences, G_Q05g and F_Q02d, on the *ICT at work* and *influence* scales respectively. In general, any subscale flagged for exclusion was removed from the set of reported scales; however, there were two scales that were retained in spite of the exclusion flag. The *writing at home* scale had a low reliability, but it was retained to maintain consistency with the reporting of at home/at work variables. The *readiness to learn* scale did have notable between country differences, but it was also fairly reliable (0.85).

The following scales were retained:

- ICT at home (7 items)
- ICT at work (7 items)
- Influence (7 items)
- Learning at work (3 items)
- Numeracy at home (6 items)
- Numeracy at work (6 items)
- Planning (3 items)
- Reading at home (8 items)
- Reading at work (8 items)
- Readiness to learn (6 items)
- Task discretion (4 items)
- Writing at home (4 items)
- Writing at work (4 items)

20.5.7 Summary of weighted likelihood estimates

For each of the retained subscales, there are a notable number of examinees with the lowest possible score (these are not always the same examinees). For the remainder of the examinees, the distributions of WLE are unimodal and appear to be approximately normal. Figure 20.2 illustrates this pattern for two of the subscales, *ICT at work* and *numeracy at work*. In these examples, if examinees do not use ICT and/or numeracy at work, there is little justification for providing scores on these subscales. As such, a decision was made to recode these values for each scale as missing. This decision is grounded in the fact that for many self-report scales of activities, zero-inflated counts are found for those respondents for which the questions are not applicable (e.g. Goodman, 1975; Dayton & Macready, 1980; Yamamoto, 1989; Loeys et al., 2012). Also note that for *ICT at home*, no such phenomenon is found as respondents were not given the *ICT at home* questions if they responded that they never used a computer before. Table 20.5 presents the means and standard deviations of each subscale when these estimates are included/excluded. By excluding these estimates in the computation of the descriptive statistics, the mean for each scale increases by 0.07 to 0.79 logits (mean = 0.33), while the standard deviations decrease by 0.06 to 0.78 logits (mean = -0.45).

Figure 20.2: WLE distributions

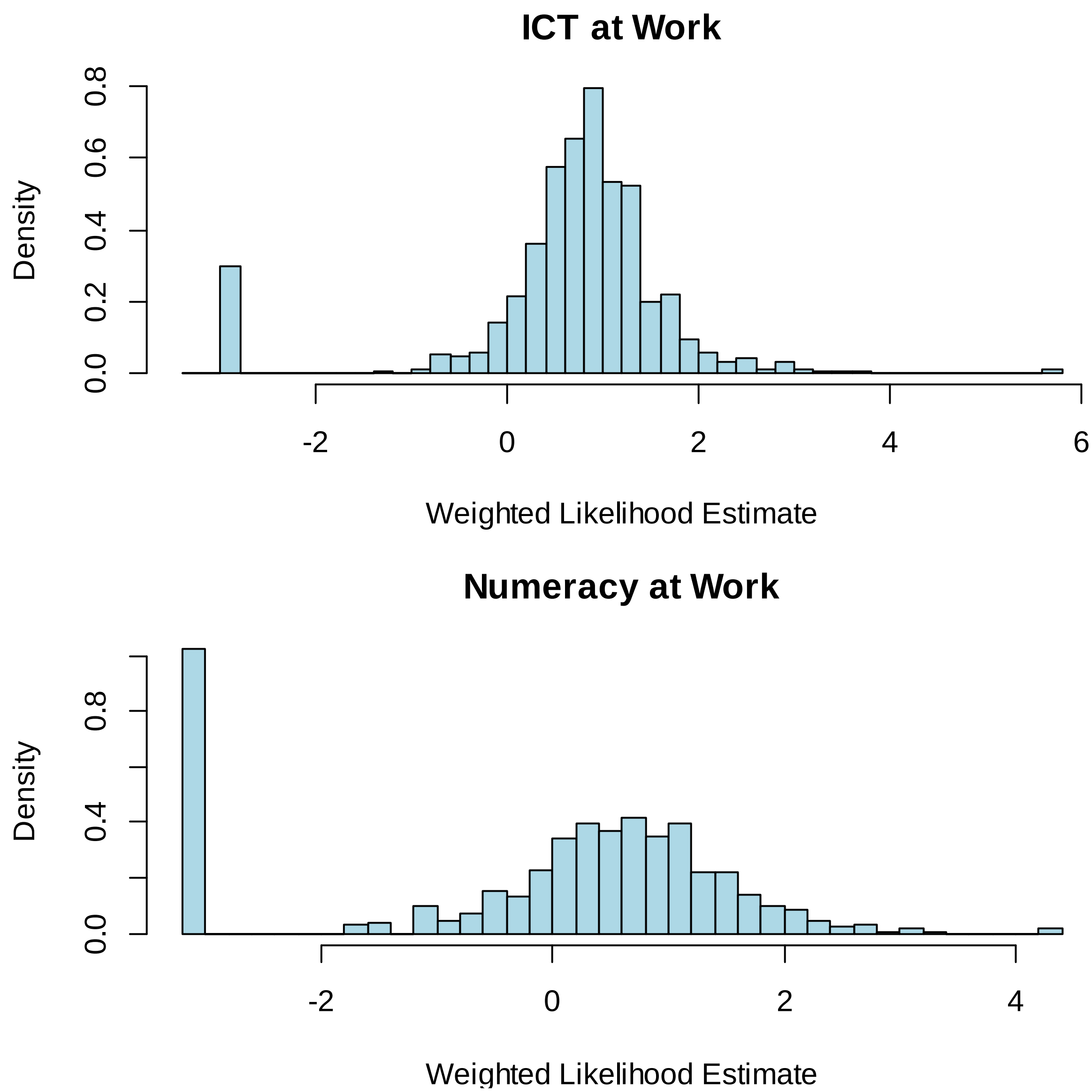


Table 20.5: WLE descriptive statistics

	Lowest Values Included		Lowest Values Excluded	
	Mean	SD	Mean	SD
ICT at Home	0.08	1.25	0.15	1.02
ICT at Work	0.15	2.00	0.54	1.37
Influence	-0.56	2.07	-0.06	1.55
Learning at Work	0.47	1.02	0.56	0.84
Numeracy at Home	-0.51	1.18	-0.21	0.76
Numeracy at Work	-0.97	1.85	-0.18	1.07
Planning	0.49	1.83	0.96	1.30
Reading at Home	0.05	1.53	0.15	1.24
Reading at Work	-0.67	2.42	-0.17	1.56
Readiness to Learn	0.81	0.58	0.83	0.52
Task Discretion	0.44	0.84	0.54	0.64
Writing at Home	-0.71	1.08	-0.43	0.69
Writing at Work	-1.00	2.01	-0.29	1.25

20.6 Item level nonresponse rates

Table 20.6 summarizes the mean proportion of missing responses across countries for all items on each scale. The standard deviation in the table indicates the variability in missing responses across countries. For most of the scales, approximately 20% to 30% of the responses were missing. With the exception of *reading at home* (SD = 19%), the variability in missing responses is fairly consistent across scales.

Table 20.6: Proportion of missing responses by scale

	Mean	SD
ICT at Home	0.01	0.03
ICT at Work	0.25	0.07
Influence	0.26	0.08
Learning at Work	0.34	0.09
Numeracy at Home	0.01	0.03
Numeracy at Work	0.25	0.07
Planning	0.26	0.07
Reading at Home	0.42	0.19
Reading at Work	0.26	0.07
Readiness to Learn	0.20	0.10
Task Discretion	0.33	0.08
Writing at Home	0.01	0.03
Writing at Work	0.25	0.07

References

Cramer, J. S. (2004). The early origins of the logit model. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 35*(4), 613-626.

Dayton, C. M., & Macready, G. B. (1980). A scaling model with response errors and intrinsically unscalable respondents. *Psychometrika, 45,* 343-356.

Goodman, L. A. (1975). A new model for scaling response patterns: An application of the quasi-independence concept. *Journal of the American Statistical Association, 70,* 755-768.

Eisinga, R., Te Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach or Spearman-Brown? *International Journal of Public Health. 58*(4): 637-642. doi:10.1007/s00038-012-0416-3

Loeys, T., Moerkerke, B., De Smet, O., & Buysse, A. (2012). The analysis of zero-inflated count data: Beyond zero-inflated Poisson regression. *British Journal of Mathematical and Statistical Psychology, 65,* 163-180. doi: 10.1111/j.2044-8317.2011.02031.x

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47,* 149-174.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, 16,* 159-176.

Yamamoto, K. (1989). *HYBRID model of IRT and latent class models.* (Research Report No. RR-89-41). Princeton, NJ: Educational Testing Service.

Chapter 21: PIAAC Proficiency Scales

Claudia Tamassia and Mary Louise Lennon, ETS

21.1 Introduction

To adequately measure the skills of adults with differing educational backgrounds and life experiences, PIAAC includes tasks that range from very easy to very challenging. As described in Chapter 2, these tasks were developed to measure the range of skills and abilities defined in the frameworks for the three assessment domains – literacy, numeracy and problem solving in technology-rich environments (PSTRE). Results from the assessment are reported along three proficiency scales, each ranging from 0 to 500 with tasks at the lower end of the scale being easier than those at the higher end.

Reporting that one task falls at 215 on a scale while another falls at 345 provides some information – namely that the first task is easier than the second – but it does not tell us much about the underlying skills and knowledge each requires. To provide a richer report of the PIAAC results, described proficiency scales were developed for each of the domains, describing what performance at various points along those scales means. To create these described proficiency scales, the expert groups in each domain met with psychometricians and test developers to review the Main Study data, look at the tasks as they were distributed along the 500-point scales, and articulate how the requisite skills and knowledge to complete those tasks progressively increased along the scale. Defining clusters of tasks which required similar skills and knowledge and differentiating them from other clusters which were more or less difficult allowed the experts to define the levels of performance along the proficiency scale for each of the PIAAC domains.

21.2 Defining the proficiency levels

The IRT scaling procedures used in PIAAC constitute a statistical solution to the challenge of establishing a scale for a set of tasks with an order of difficulty that is essentially the same for everyone.

First, the response data collected from each participating country was used to estimate item parameters for each scale using a particular IRT model. In PIAAC, a two-parameter model was used that models the probability of a response based on the difficulty of an item and how well it discriminates, in combination with the person's ability or proficiency. This information was summarized in the form of item characteristic curves which show the probability of successfully completing an item at a given level of ability. Next, item parameters along with other information were used to estimate the ability distributions for each participating country along a scale with an overall mean and standard deviation. This scale can then be used to compare the overall performance of countries or subgroups within a country. It can also be used to compare performance along the scale based on statistical criteria such as percentiles.

The IRT analysis summarizes how well the sample of individuals who responded to the pool of tasks performed. The tasks in this pool constitute a sample of the universe or "population" of tasks representing the construct that is measured (in the case of PIAAC, literacy, numeracy and PSTRE as defined by the relevant framework documents). Thus, the goal is to make inferences concerning the proficiency of respondents with respect to the population of tasks that represent the construct – that is, to make inferences about how well respondents performed on items used in the assessment as well as items having similar characteristics that also represent the construct but were not included in this particular assessment. As the items used in the survey represent a sample of tasks, it is important that any description of skills closely align to the framework used to define and construct them.

The use of IRT makes it possible not only to summarize results for various subpopulations of adults but also determine the relative difficulty of the tasks. In other words, just as individuals receive a specific score along a scale according to their performance on the assessment tasks, each task receives a specific value on a scale according to its difficulty, as determined by the performance of adults across the various countries that participated in the assessment (Kirsch et al., 2002). As tasks used in PIAAC vary widely in terms of task requirements and levels of complexity, it is possible to capture the range of difficulty of task through an item map which places items along a scale based on a selected response probability.[1]

Test items do not discriminate perfectly and each person has a chance (however small) of responding correctly to any given item. Consequently, a value representing the probability of correctly responding to an item must be selected in order to place an item on a proficiency scale. In theory, any value greater than zero and less than one can be chosen to place items on a proficiency scale, and a range of RP values are used in large-scale assessments. A value of 0.62 is used in PISA (OECD, 2009). Trends in International Mathematics and Science Study (TIMSS) uses different values for constructed responses (0.50) and multiple choice items (0.65) (TIMSS, 2007). The US National Assessment of Educational Progress (NAEP) uses an RP of 0.74 for multiple-choice items and 0.65 for open-ended items (National Center for Education Statistics, 2011). The IALS and ALL surveys used an RP of 0.80. The US National Assessment of Adult Literacy (NAAL) used an RP of 0.80 in reporting its 1992 survey and 0.67 in reporting results from its 2002 survey (Hauser, Edler, Koenig, & Elliott, 2005).

In PIAAC, the OECD Secretariat and participating countries agreed on an RP value of 0.67, similar to the approach used in PISA, to ensure that the description of what it means to be performing at a particular level of proficiency is consistent between the two surveys. There are potential risks for the credibility of both studies if being at a particular level of proficiency meant something different in each survey. While the RP value used in PIAAC and PISA will not be identical,[2] the interpretation of what it means to be at a level of proficiency will be the same.

[1] The RP section of this chapter was based on a PIAAC BPC document, Proficiency Levels in PIAAC [Doc. Ref.: COM/DELSA/EDU/PIAAC(2011)14], and written by Irwin Kirsch and Kentaro Yamamoto.

[2] This is a result of the different widths of the proficiency bands used.

Within any given scale, except for those at the lowest level, a person would be expected to pass a test made up of items from the level at which he or she performed. For example, using RP67, a person at the bottom of Level 3 on the literacy scale would be expected to successfully complete items of Level 3 difficulty approximately 50 percent of the time, a person at the top of the level would be expected get such items correct around 80 percent of the time, and a person at the middle of the level would do so 67 percent of the time. The probability of success on Level 3 items of persons at the top, bottom and middle of Level 3 based on RP80 is approximately 60, 80 and 90 percent, respectively. It is important to note that for both RP values, a person at the middle of a level would be likely to get most items at a lower level correct as well as a reasonable proportion of items at the next highest level. It is also important to note that the selection of a response probability is independent from the estimation of both item parameters and ability. The choice of an RP value has no impact on either the statistical characteristics of the items or the estimation of ability along the scale. In addition, the precision of measurement along a scale is not affected by the RP value. The same items define the underlying scale regardless of which RP value is selected.

As RP80 was used in IALS and ALL, in order to ensure that countries that wish to do so can map the change from RP80 to RP67, the OECD Secretariat provided item maps for literacy and numeracy under both the PISA approach (RP67) and the RP80 assumption in an appendix to the international report.

21.3 Interpreting the proficiency levels

As explained in the previous section, the proficiency scales range from 0 to 500 and are designed so the scores represent degrees of proficiency in a particular aspect of the domain. There are easier and harder tasks for each proficiency scale.[3] Each scale is divided into proficiency levels based on the knowledge and skills required to complete the tasks within those levels.

The purpose of described proficiency scales is to facilitate the interpretation of the scores assigned to respondents. That is, respondents at a particular level not only demonstrate knowledge and skills associated with that level but also the proficiencies required at lower levels. Thus, respondents scoring at Level 2 are also proficient at Level 1, with all respondents expected to answer at least half of the items at that level correctly.

The PIAAC proficiency scales and item descriptions were part of the work done by the PIAAC Expert Groups in December 2012 and January 2013. For a complete list of experts in these groups, please see Appendix 6.

21.3.1 Literacy

As described in Chapter 2 of this report, the PIAAC literacy items were developed and selected to represent three major aspects of processing continuous and noncontinuous texts and documents: accessing and identifying, integrating and interpreting, and reflecting on and evaluating information.

- *Access and identify* tasks require the reader to locate information in a text or document. While some tasks can be relatively straightforward because the information requested in

[3] See Appendix 1 for the complete list of Main Study PIAAC items in each domain organized by difficulty.

the question matches clearly with information that is easily located in the text, not all tasks in this category are necessarily easy. Inferences may need to be made and rhetorical understanding may be required.

- *Integrate and interpret* tasks require the reader to relate different parts of the text to each other. Requiring respondents to compare and contrast, understand problems and solutions, and identify cause/effect relationships are examples of this task type. These relationships may be explicitly signaled (e.g., the text states that "the cause of X is Y") or may require the reader to make inferences. The text components to be related may be contiguous and therefore easier to locate and integrate or may be found in different paragraphs in the same text or in separate documents.

- *Evaluate and reflect* tasks require the reader to draw on knowledge, ideas or values external to the text. The reader must assess the relevance, credibility, argumentation and truthfulness of the information presented in the text within a context of information that is not present in the text. The reader may also evaluate the purposefulness, register, structure or reader awareness of the text, or the success with which the author uses evidence and language to argue or persuade. Tasks of this type were judged to be particularly important to include in the context of PIAAC's digital texts, where it is readers must be alert to a text's accuracy, reliability and timeliness.

The PIAAC literacy framework defined features of stimulus texts and tasks that were anticipated to impact the difficulty of tasks included in the assessment.[4] These included the following:

- transparency of information in the text as it relates to the presented task or question
- degree of complexity necessary to make required inferences
- semantic and syntactic complexity of the text and/or question
- amount of text that must be processed
- prominence of needed information in the text
- competing information in the text
- text features that facilitate or hinder understanding relationships among parts of the text

The literacy proficiency scale is defined in terms of six levels. In all, the literacy scale includes 58 tasks with that ranged in difficulty from an RP67 of 75 to 376. Those tasks are distributed by level as follows:

- Below Level 1 (1 – 175): 4 tasks
- Level 1 (176 – 225): 3 tasks

[4] For the full text of the PIAAC Literacy Framework, see Chapter 3 of OECD (2012).

- Level 2 (226 – 275): 15 tasks
- Level 3 (276 – 325): 24 tasks
- Level 4 (326 – 375): 11 tasks
- Level 5 (376 – 500): 1 task

Each of the six proficiency levels is defined below and one or more representative tasks are described to illustrate the key information-processing skills at each level.

Literacy Below Level 1	**0 to 175**
The tasks at this level require the respondent to read brief texts on familiar topics to locate a single piece of specific information. Only basic vocabulary knowledge is required, and the reader is not required to understand the structure of sentences or paragraphs or make use of other text features. There is seldom any competing information in the text and the requested information is identical in form to information in the question or directive. While the texts can be continuous, the information can be located as if the text were noncontinuous. Tasks below Level 1 do not make use of any features specific to digital texts.	

SGIH (C301AC05)
Difficulty: 75
In this task, respondents are asked to identify a telephone number in a very short advertisement. The question explicitly refers to literal information in a simple text with little competing information. The information is prominently located on a single line in the advertisement, labeled by an abbreviation for the word “telephone.” These features of the text and question combine to make this the easiest task on the PIAAC literacy scale.

Election Results (C302BC02)
Difficulty: 162
Respondents are asked to use a notice providing results from a union election to identify the candidate with the fewest number of votes. Although the notice contains several paragraphs of information, the respondent only needs to use a very short table with three numbers and associated names within the text to answer the question. The key word (“votes”) appears in both the prompt and the text making the relevant information very transparent. There is no competing information as the word “votes” appears nowhere else in the text. To locate the answer, the respondent needs to compare the three numbers (the word “fewest” in the prompt indicates the answer will involve a number), and once that is determined, locate the name associated with that number.

Literacy Level 1	**176 to225**
Most of the tasks at this level require the respondent to read relatively short digital or print continuous, noncontinuous or mixed texts to locate a single piece of information which is identical to or synonymous with the information given in the question or directive. Some tasks may require the respondent to enter personal information into a document, in the case of some noncontinuous texts. Little, if any, competing information is present. Some tasks may require simple cycling through more than one piece of information. Knowledge and skill in recognizing basic vocabulary, evaluating the meaning of sentences, and reading of paragraph text is expected.	

Dutch Women (C311B701)
Difficulty: 201
This task asks the respondent to find the percentage of women in the teaching profession in Greece based on a graphically presented table showing that information for 10 countries. There is a single instance of the word "Greece" in the stimulus and a single instance of a percentage associated with that word, making the task relatively simple. There are other percentages in the text that might serve as distractors or cause the respondent to misread the table, which makes this more difficult than the Below Level 1 tasks, but the explicit connection between the question wording and information in the stimulus makes this a relatively simple task.

Generic Medicine (C309A321)
Difficulty: 219
This stimulus consists of a short newspaper article focusing on the limited use of generic medicines in Switzerland. The article includes a simple two-column table showing the market share for generic medications in 15 countries. The Level 1 item associated with this stimulus asks the respondent to identify the number of countries where generic medicines account for more than 10% of drug sales. While the phrase "drug sales" does not appear in the text, the only place a list of countries and percentages appears is in the table included in the article. The phrase "market share" is in the title of this table and might be regarded as a synonym for "drug sales," but most respondents would not need this additional information. The respondent's task is then to simply count the number of percentages that are greater than 10%, a task made simpler as the percentages are ordered from large to small.

Literacy Level 2	**226 to 275**
At this level, the complexity of text increases. The medium of texts may be digital or printed, and texts may comprise continuous, noncontinuous or mixed types. Tasks in this level require respondents to make matches between the text and information, and may require paraphrase or low-level inferences. Some competing pieces of information may be present. Some tasks require the respondent to • cycle through or integrate two or more pieces of information based on criteria, • compare and contrast or reason about information requested in the question, or • navigate within digital texts to access and identify information from various parts of a document.	

Lakeside Fun Run (C322P002)
Difficulty: 240
This unit is based on a Web page with information about a community relay race and walking event. The tasks associated with the unit require some understanding of Web conventions. This task, the easiest in the unit, asks respondents to identify the link they would use to find the phone number for one of the event organizers. The correct response, a link labeled "Contact Us," is one of several on the home page of this digital text. While using this link might be apparent to respondents familiar with Web-based texts, less familiar respondents need to make some inferences in order to know where to navigate to find the information.

Generic Medicines (C310A406)
Difficulty: 272
This task uses the same stimulus as that described in Level 1 above but requires the respondent to use the text of the newspaper article. Here the respondent is asked to identify two reasons given in the text for the limited use of generic medicines. Previous research has shown that tasks requiring multiple responses tend to be more difficult as respondents must search through the text more than once. While the reasons are explicitly stated in the text, they are not specifically labeled as reasons. Respondents must make an inference based on a semantic cue in the text – the single word "Why?" which signals that reasons will follow. There are other instances of "reasons" in the text (such as why generic medicines are less expensive, signaled by the explicit "because") that might serve as distractors for less able respondents.

Literacy Level 3 **276 to 325**
Texts at this level are often dense or lengthy, including continuous, noncontinuous, mixed or multiple pages. Understanding text and rhetorical structures become more central to successfully completing tasks, especially in navigation of complex digital texts. Tasks require the respondent to identify, interpret or evaluate one or more pieces of information and often require varying levels of inferencing. Many tasks require the respondent construct meaning across larger chunks of text or perform multistep operations in order to identify and formulate responses. Often tasks also demand that the respondent disregard irrelevant or inappropriate text content to answer accurately. Competing information is often present, but it is not more prominent than the correct information.

Lakeside Fun Run (C322P001)
Difficulty: 283
This question in the "Lakeside Fun Run" unit asks the respondent to identify information in the Web page that explains how this year's race differs from last year's. Not only does the task require the respondent to understand a contrast – a more difficult semantic construct – but the contrast is only indirectly signaled in the text, which says, "The popular walk will continue, but this year…"

Lakeside Fun Run (C322P004)
Difficulty: 293
A more difficult task from the "Lakeside Fun Run" task requires the respondent to understand a common convention in digital texts – a FAQ (frequently asked questions) link – and be able to

use it to navigate through the text. The respondent is asked to identify the date by which a race participants must notify organizers they want to change their race distances. In order to find the requested information, the respondent must click on the FAQ link on the home page. Once the respondent has successfully navigated to the FAQ page, the information on the page is relatively easy to find, as there is a near synonymous match between the task statement and the text.

Literacy Level 4	**326 to 375**
Tasks at this level often require respondents to perform multiple-step operations to integrate, interpret, or synthesize information from complex or lengthy continuous, noncontinuous, mixed, or multiple type texts. Complex inferences and application of background knowledge may be needed to perform successfully. Many tasks require identifying and understanding one or more specific, noncentral ideas in the text in order to interpret or evaluate subtle evidence claim or persuasive discourse relationships. Conditional information is frequently present in tasks at this level and must be taken into consideration by the respondent. Competing information is present and sometimes seemingly as prominent as correct information.	

Library Search (C323P004)
Difficulty: 329
The stimulus for this unit consists of two pages from a library website listing results for a search on "genetically modified food." This task asks the reader to find two books that argue against genetically modified foods, requiring the respondent to examine the brief descriptions of all the books and decide which best meet that criterion. The respondent must scroll through the full list, using both pages on the website, to make inferences and compare the descriptions in the 10 entries. As the task asks for two books, the respondent must cycle through the text twice to locate both responses.

Library Search (C323P002)
Difficulty: 348
The same "Library Search" unit includes another example of a Level 4 task that is harder than the task above. The task asks the respondent to find the single book that suggests that the claims both for and against genetically modified foods are unreliable. The information in the text that the respondent uses to find the answer is "manufactured propaganda," which the respondent has to infer is meant to be synonymous with the word "unreliable" that is in the prompt. The task requires the careful respondent to examine all the entries.

Literacy Level 5	**376 to 500**
At this level, tasks may require the respondent to search for and integrate information across multiple, dense texts; construct syntheses of similar and contrasting ideas or points of view; or evaluate evidence-based arguments. Application and evaluation of logical and conceptual models of ideas may be required to accomplish tasks. Evaluating reliability of evidentiary sources and selecting key information is frequently a key requirement. Tasks often require respondents to be aware of subtle, rhetorical cues and to make high-level inferences or use specialized background knowledge.	

Library Search (C323P005)
Difficulty: 376
One of the most difficult literacy tasks in PIAAC is also associated with the "Library Search" unit. The respondent is asked to identify the book likely to be least useful in providing more information about genetically modified food. As mentioned in the framework, negative phasing is more complex than affirmative, so evaluating the 10 books in terms of which is *least* useful for the defined purpose is expected to be difficult. The fact that the correct selection is located at the end of the second page of results also increases the difficulty of the task. The respondent must read and evaluate each of the choices in order to make a correct selection.

21.3.2 Numeracy

The PIAAC numeracy framework includes a definition of the domain as well as a description of numerate behavior.[5] Numeracy tasks were developed to cover a range of difficulty as a result of combining variables that include:

- the kind and degree of interpretation and reflection required by the problem,
- the kind of representation skills required,
- the kind and level of mathematical skill required (e.g., single-step vs. multistep problems, or more advanced mathematical knowledge, complex decision making, and problem-solving and modeling skills),
- the kind and degree of mathematical argumentation required,
- the degree of familiarity with the context, and
- the extent to which tasks require reproduction of known procedures and steps or present novel situations requiring nonroutine and perhaps more creative responses.

The numeracy proficiency scale is defined in terms of six levels and includes 56 tasks with difficulty values ranging from 129 to 375. Based on RP67, these tasks are distributed by level as follows:

- Below Level 1 (1 – 175): 3 tasks
- Level 1 (176 – 225): 6 tasks
- Level 2 (226 – 275): 21 tasks
- Level 3 (276 – 325): 20 tasks
- Level 4 (326 – 375): 6 tasks

[5] For the full text of the PIAAC Numeracy Framework, see Chapter 4 of OECD (2012).

- Level 5 (376 – 500): 0 tasks

Each of the six proficiency levels is defined below and one or more representative tasks are described to illustrate the key skills and knowledge at each level.

Numeracy Below Level 1	**0 to 175**
Tasks at this level are set in concrete, familiar contexts where the mathematical content is explicit with little or no text or distractors and that require only simple processes such as counting, sorting, performing basic arithmetic operations with whole numbers or money, or recognizing common spatial representations.	

Bottles (C601AC06)
Difficulty: 129

The easiest task on the numeracy scale, with difficulty level of 129, requires respondents to look at a photograph containing two cases of water bottles. They are asked to find the total number of bottles in the two full cases being shown. Part of what makes this task easy is that content is drawn from everyday life and objects of this kind are relatively familiar to most people. Second, what respondents are asked to do is apparent and explicit – this task uses a photograph depicting concrete objects and containing no text to be read. A third contributing factor is that respondents can approach the task in a variety of ways that differ in sophistication, such as by multiplying rows and columns, but also by simple counting. This task requires that adults make a conjecture using spatial visualization because the full set of bottles in the lower case is not visible, but as can be seen from the low difficulty level of the task, this feature did not present a problem for the vast majority of adults in participating countries.

Numeracy Level 1	**176 to225**
Tasks in this level require the respondent to carry out basic mathematical processes in common, concrete contexts where the mathematical content is explicit with little text and minimal distractors. Tasks usually require simple one-step or two-step processes involving, for example, performing basic arithmetic operations; understanding simple percents such as 50%; or locating, identifying and using elements of simple or common graphical or spatial representations.	

Tea Candles (C615A602)
Difficulty: 221

An example of a Level 1 task is Tea Candles Q1. The stimulus for this item consists of a photo of a box containing tea light candles. The packaging identifies the product (tea light candles), the number of candles in the box (105 candles) and its weight. While the packaging partially covers the top layer of candles, it can be seen that the candles are packed in five rows of seven candles each. The instructions inform the respondent that there are 105 candles in a box and asks him or her to calculate how many layers of tea candles are packed in the box.

Numeracy Level 2	**226 to 275**
Tasks in this level require the respondent to identify and act upon mathematical information and ideas embedded in a range of common contexts where the mathematical content is fairly explicit or visual with relatively few distractors. Tasks tend to require the application of two or more steps or processes involving, for example, calculation with whole numbers and common decimals, percents and fractions; simple measurement and spatial representation; estimation; and interpretation of relatively simple data and statistics in texts, tables and graphs.	

Gas Gauge (C604A505)
Difficulty: 228

This is a somewhat more complex numeracy task falling in the lower end of Level 2. A gauge is presented that has three lines or ticks on it: one showing an "F," one showing an "E" and one in the middle of the others. A line on the gauge, representing the gauge's needle, shows a level that is roughly halfway between the middle tick and the tick indicating "F," suggesting that the tank is about three-quarters full. The task states that the tank holds 48 gallons and asks the respondent to determine "how many gallons remain in the tank." This task is drawn from an everyday context and requires an adult to interpret a display that conveys quantitative information but carries virtually no text or numbers. No mathematical information is present other than what is given in the question. What makes this task more difficult than the previous ones is that adults must first estimate the level of gas remaining in the tank by converting the placement of the needle to a fraction. Then they need to determine how many gallons this represents from the 48-gallon capacity stated in the question. Thus, this task requires adults to apply multiple operations or procedures to arrive at a correct response without specifying what the operations may be. Nonetheless, this task, like many everyday numeracy tasks, does not require an exact computation but allows an approximation that should fall within reasonable boundaries.

Cooper Test (C601AC06)
Difficulty: 234

This Level 2 item engages the respondent with moderately complex tables of numerical and textual data relating to a common measure of physical fitness – the Cooper Test – from which they have to read off the level of fitness of a 43-year-old male who runs 1,100 meters in 12 minutes. This task is drawn from everyday life and involves interpreting the headings and numerical information in the table correctly in order to locate the 40-49 age table row and the appropriate cell in this row for a male who runs 1,100 meters in the requisite 12 minutes. There is no calculation involved, but number bands for both age and distance need to be understood. However, it is a type of task many adults, particularly those who use the Internet regularly, would have experienced.

Numeracy Level 3 **276 to 325**

Tasks in this level require the respondent to understand mathematical information which may be less explicit, embedded in contexts that are not always familiar, and represented in more complex ways. Tasks require several steps and may involve the choice of problem-solving strategies and relevant processes. Tasks tend to require the application of, for example, number sense and spatial sense; recognizing and working with mathematical relationships, patterns, and proportions expressed in verbal or numerical form; and interpretation and basic analysis of data and statistics in texts, tables and graphs.

Tiles (C619A609)
Difficulty: 282

This Level 3 item presents the respondent with a plan of a kitchen floor to be tiled with nine of the proposed square tiles placed in a corner, with the plan drawn on a squared grid. It asks the respondent to use this information to find out how many tiles are needed to cover the entire floor. The task is a familiar one drawn from everyday life and, using the most obvious method an adult would choose, would require several operations to arrive at the correct answer. First, the area in terms of the number of larger grid squares in the kitchen floor plan is calculated by counting or otherwise. Then the number of tiles in each larger square is calculated by counting or multiplication. The last step involves multiplying the number of larger squares by the number of tiles per larger square to get the total number of tiles required to cover the kitchen floor. Respondents need to use their spatial reasoning ability in organizing the information in the first two steps in this task. The task could also be done using a combination of spatial visualization and counting all the small squares (tiles), but this method would be more prone to error.

Orchestra Tickets (C664P001)
Difficulty: 307

This task has a difficulty around the middle of Level 3. It presents the respondent with a table of numerical data on ticket price categories for single and multiple events (Season Ticket). The respondent has to discern the pattern in the data and identify the formula, probably in verbal or numerical terms (e.g., multiply by 4½), for calculating the cost of a season ticket from the cost of a single ticket for different seating categories to an event, and use it to calculate the cost of a season ticket for a new entry category – a student season ticket. The task requires adults to use a range of reasoning strategies, including algebraic reasoning (i.e., reasoning with variables and generalizing from specific values) and computational procedures.

Numeracy Level 4 **326 to 375**

Tasks in this level require the respondent to understand a broad range of mathematical information that may be complex, abstract or embedded in unfamiliar contexts. These tasks involve undertaking multiple steps and choosing relevant problem-solving strategies and processes. Tasks tend to require analysis and more complex reasoning about, for example, quantities and data; statistics and chance; spatial relationships; change; proportions; and formulas. Tasks in this level may also require comprehending arguments or communicating well-reasoned explanations for answers or choices.

Cooper Test (C665P002)
Difficulty: 326

This task is based on the same stimulus as the Level 2 task described above but was considerably more difficult for adults in participating countries. It requires respondents to go beyond interpreting the information in the tables to calculate the percent increase needed in the distance run by a female in 12 minutes for her fitness level to be in the "Good" category. To arrive at a correct response, respondents have to locate the "Good" band for a 27-year-old female and use the difference between the runner's current 12-minute distance and the minimum distance for the "Good" band to calculate the percent increase in distance run by her to qualify for that band. There is considerable use of reasoning and knowledge and understanding of percentages in carrying out this task.

Compound Interest (P610A515)
Difficulty: 348

This is the third most difficult task in the PIAAC numeracy assessment. It presents respondents with an advertisement claiming it is possible for an investor to double an amount invested in seven years, based on a 10 percent fixed interest rate each year. Adults are asked if it is possible to double $1,000 invested at this rate after seven years and have to support their answer with their calculations. A range of responses was accepted as correct as long as a reasonable justification was provided, with relevant computations. Respondents were free to perform the calculation any way they wanted, but they could use a "financial hint," which accompanied the advertisement and presented a formula for estimating the worth of an investment after a specified number of years. Those who used the formula had to enter information stated in the text into variables in the formula (principal, interest rate and time period) and then perform the needed computations and compare the result to the expected amount if $1,000 is doubled. All respondents could use a handheld calculator provided as part of the assessment.

This task proved difficult because it involved percents, and the computation, with or without the formula, required the integration of several steps and several types of operations. Performing the computations without the formula required understanding of compound interest procedures. This task required adults to use a range of reasoning strategies, including algebraic reasoning and informal or invented procedures. It also required the use of formal mathematical information and deeper understanding of nonroutine computational procedures, all of which may not be familiar or accessible to many adults.

Numeracy Level 5	**376 to 500**

Tasks in this level require the respondent to understand complex representations and abstract and formal mathematical and statistical ideas, possibly embedded in complex texts. Respondents may have to integrate multiple types of mathematical information where considerable translation or interpretation is required; draw inferences; develop or work with mathematical arguments or models; and justify, evaluate and critically reflect upon solutions or choices.

21.3.3 Problem solving in technology-rich environments

The PSTRE domain is organized around three core dimensions: the cognitive strategies and processes a person uses to solve a problem, the tasks or problem statements that trigger and condition problem solving, and the technologies through which the problem solving is conducted. Variations within and across all of those dimensions were expected to contribute to the overall difficulty of the problems presented in the PIAAC assessment. For example, a problem is likely to be more complex if it is ill-defined as opposed to explicitly stated, if it requires complex problem solving strategies such as defining goals and resolving impasses, and/or if it requires the use of multiple technology environments (e.g., respondents must utilize both emails and spreadsheets).

In order to explain how proficiency can be affected by the three dimensions of PSTRE, the problem-solving proficiency scale was divided into three levels as shown below. In this section, we describe the essential features of tasks at each of these three levels.

Table 21.1: Technology, task and cognitive characteristics of problems at each of three main levels of proficiency

Level	Technology features	Task features	Cognitive processes
Level 1	• Generic applications • Little or no navigation required • Relevant information is directly available • Use of facilitating tools not required	• Few steps • Single operators	• Reach a given goal • Apply explicit criteria • Minimal monitoring demands • Simple relevance match • Categorical reasoning • No integrate or transformation
Level 2	• Both generic and novel applications (e.g., Web-based services) • Some navigation required to acquire information or perform actions • Use of tools facilitates operations	• Multiple steps • Multiple operators	• Goal may need to be defined • Apply explicit criteria • Generally higher monitoring demands • Generally involves resolving impasses • Some evaluation of relevance • Some integrate or transformation • Inferential reasoning
Level 3	• Generic and novel applications • Some navigation required to acquire information or perform actions • Use of tools required to efficiently solve the problem	• Multiple steps • Multiple operators	• Goal may need to be defined • Establish and apply criteria • Generally high monitoring • High inferential reasoning and integration • Evaluate relevance and reliability • Generally involves resolving impasses

The proficiency levels of PSTRE are defined as follows:

PSTRE Below Level 1 **0 to240**
Tasks are based on well-defined problems involving the use of only one function within a generic interface to meet one explicit criterion without any categorical, inferential reasoning or transforming of information. Few steps are required and no subgoal has to be generated.

Though the current set of tasks included very simple problems, none of those fell within the Below Level 1 category. The simplest item on the assessment had an RP67 of 268. The expert group did, however, consider the characteristics of tasks that might fall at this level. Based on the PSTRE framework (OECD, 2012), such problems would have the following characteristics. They would be well-defined problems involving the use of only one function on a generic interface to meet one explicit criterion without any categorical, inferential reasoning or transforming of information. Few steps would be required and no subgoal would have to be generated. PSTRE problems at this level would still differ from simple ICT literacy in that the goal would extend beyond the mere use of ICT functions and commands. Thus, respondents would still need to implement a set of actions aimed at solving the problem through the use of technology.

It should be noted that more than a quarter of the PIAAC participants were excluded from the PSTRE survey because they reported no prior experience using computers, they were not willing to take the survey on a computer, or they were not able to demonstrate the basic ICT skills required to complete the assessment such as clicking, highlighting and simple typing. This proportion is likely to decrease in future surveys, as more and more people become familiar with using computers and other digital devices such as smartphones and tablets. It is likely that future assessment would include a larger percentage of the total population, most of which would likely display modest levels of proficiency. Therefore, in future assessments it will become increasingly important to include easier tasks to better describe in more detail the lower end of the proficiency scale

PSTRE Level 1 **241 to 290**
At this level, tasks typically require the use of widely available and familiar technology applications, such as email software or a Web browser. There is little or no navigation required to access the information or commands required to solve the problem. The problem may be solved regardless of one’s awareness and use of specific tools and functions (e.g., a sort function). The task involves few steps and a minimal number of operators. At a cognitive level, the person can readily infer the goal from the task statement; problem resolution requires one to apply explicit criteria; there are few monitoring demands (e.g., the person does not have to check whether he or she has used the adequate procedure or made progress toward the solution). Identifying contents and operators can be done through simple match; only simple forms of reasoning, for example, assigning items to categories are required. There is no need to contrast or integrate information.

Party Invitations (U01A)
Difficulty: 286

This task presents a problem where respondents are asked to organize a set of email responses they had received in response to a party invitation. The necessary folders are present in the email environment; respondents need to sort a set of emails into those existing folders. The email interface is presented with five emails in an inbox and the respondent is asked to organize the responses to keep track of who can and cannot attend the party. In terms of the three PSTRE dimensions, the item requires the respondent to categorize a small number of messages in an email application in existing folders according to a single criterion. This is typical of a Level 1 item because the goal is explicitly stated in operational terms, the task is performed in a single environment, and it can be solved in a relatively small number of steps using a restricted range of operators. Thus, the task does not require the user to learn a novel environment, nor does it necessitate a significant amount of monitoring across a large number of actions.

PSTRE Level 2 **291 to 340**
At this level, tasks typically require the use of both generic and more specific technology applications. For instance, the person may have to make use of a novel online form. Some navigation across pages and applications is required to solve the problem. The use of tools (e.g., a sort function) can facilitate the resolution of the problem. The task may involve multiple steps and operators. In terms of cognitive processing, the problem goal may have to be defined by the person, though the criteria to be met are explicit. There are higher monitoring demands. Some unexpected outcomes or impasses may appear. The task may require evaluating the relevance of a set of items to discard distractors. Some integration and inferential reasoning may be needed.

Club Membership (U19B)
Difficulty: 296

This task consists of responding to an information request and demands locating information in a spreadsheet. Respondents must identify an undefined number of members of a biking club who meet the provided eligibility requirements to serve as club president. The information can most efficiently be located within the long spreadsheet by using a sort function. The respondent is presented with two environments: a word processor page containing information about the two conditions required for club presidents, and a database with 200 entries where the relevant information can be found. In terms of the three PSTRE dimensions, the item requires the respondent to organize large amounts of information in a multiple column spreadsheet using multiple explicit criteria and locate and mark relevant entries. This is typical of Level 2 because the task requires switching between two different applications and involves multiple steps and operators. It also requires some amount of monitoring. Making use of the available tools (e.g., the sort function) greatly facilitates the identification of the relevant entries.

PSTRE Level 3	**341 to 500**
At this level, tasks typically require the use of both generic and more specific technology applications. Some navigation across pages and applications is required to solve the problem. The use of tools (e.g., a sort function) is required to make progress toward the solution. The task may involve multiple steps and operators. In terms of cognitive processing, the problem goal may have to be defined by the person, and the criteria to be met may or may not be explicit. There are typically high monitoring demands. Unexpected outcomes and impasses are likely to occur. The task may require evaluating the relevance and the reliability of information in order to discard distractors. Integration and inferential reasoning may be needed to a large extent.	

Meeting Rooms (U02)
Difficulty: 346

This task requires respondents to check a number of email requests regarding reservations for a meeting room on a particular date and schedule those reservations based on multiple constraints (including the number of rooms available and reservations already made). Impasses due to conflicting constraints have to be resolved by initiating a new subgoal, that is, issuing a standard message to decline one of the requests. Two environments are present: an email interface with a number of emails containing the requests for meeting dates and times, and a novel Web application that allows respondents to assign rooms to meetings at certain times. Upon discovering that one of the requests cannot be accommodated, the respondent has to use a specific command on the website in order to issue a standard message declining the request. In terms of the three PSTRE dimensions, the item requires the respondent to use information from a novel Web application and several email messages, establish and apply criteria to solve a scheduling problem where an impasse must be resolved, and communicate the outcome. This is typical of Level 3 as the task involves multiple applications, a large number of steps, a built-in impasse, and requires the respondent to discover and use ad hoc commands in a novel environment. The respondent has to set up and monitor the application of a plan in order to minimize the number of conflicts. Furthermore, the respondent has to transfer information from one application (email) to another (room reservation).

21.4 Final remarks

This chapter focused on described proficiency scales, an important reporting tool that enhances the understanding of what has been measured in large-scale surveys such as PIAAC and allows policymakers and other stakeholders to better interpret survey results. Each of the PIAAC expert groups reviewed the Main Study data and analyzed the characteristics of tasks that fell along the scale for each domain, defining proficiency levels and describing the cognitive skills and knowledge required at each level.

References

Hauser, R. M., Edler, C. F. Jr., Koenig, J. A., & Elliott, S. W. (Eds.) (2005). *Measuring literacy: performance levels for adults*. Retrieved from http://www.nap.edu/catalog.php?record_id=11267

Kirsch I., de Jong, J., Lafontaine, D., McQueen, J., Mendelovits, J., & Monseur, C. (2002). *Reading for change – performance and engagement across countries.* Retrieved from Organisation for Economic Co-operation and Development website: http://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/33690904.pdf

National Center for Education Statistics. (2011). *NAEP technical documentation*. Retrieved from NCES website: http://nces.ed.gov/nationsreportcard/tdw/analysis/describing_itemmapping.asp

Organisation for Economic Co-operation and Development. (2000). *Literacy in the information age – final report of the International Adult Literacy Survey*. Retrieved from OECD website: http://www.oecd.org/edu/skills-beyond-school/41529765.pdf

Organisation for Economic Co-operation and Development. (2009). *PISA 2006 technical report.* Retrieved from OECD website: http://www.oecd.org/pisa/pisaproducts/pisa2006/42025182.pdf

Organisation for Economic Co-operation and Development. (2011). *Proficiency levels in PIAAC*, (Report No. COM/DELSA/EDU/PIAAC(2011)14). Paris, France: Author.

Organisation for Economic Co-operation and Development. (2012). *Literacy, numeracy and problem solving in technology-rich environments – framework for the OECD Survey of Adult Skills*. Paris, France: Author.

Trends in International Mathematics and Science Study (TIMSS). (2007). TIMSS 2007 Technical report. Retrieved from TIMSS website: http://timss.bc.edu/timss2007/PDF/T07_TR_Chapter13.pdf

Chapter 22: Generating Results for PIAAC

Alfred Rogers and John Barone, ETS

22.1 Data processing and analysis

The ETS data analysis systems are set up to process the PIAAC data in both SPSS file format and "flat" file ASCII text format. It was therefore imperative for both sets of data files across all countries to be perfectly synchronized with respect to the currency and content of the constituent data fields.

SPSS data files are completely self-documented, containing variable labels, data value labels and missing value definitions in addition to the data. However, many of the scaling and analytic tools used by ETS required the input data to be represented in "flat" file ASCII text rectangular format, where each data field is in the same position on every record in the file. ETS developed a procedure that extracts the data from an SPSS file into an ASCII text file and also extracts the metadata (labels, formats, missing value definitions, etc.) into a proprietary XML data dictionary file. Any program or procedure that uses the ASCII data file must first process the XML dictionary file to map the contents of the data file onto the set of variables to be analyzed or processed.

22.2 Receipt processing

When the data files were received from the IEA-hosted secure FTP site, they were unzipped and placed in a date-tagged folder before transfer to the operational folder.

Many of the data variables in the survey component were long text responses that could not be reduced to numeric codes and needed to be retained in the database for future interpretation. These responses were usually encoded in the native language of each country and could contain extended ASCII codes (Unicode) to represent certain characters. When placed in an ASCII file, these codes corrupt the rectangular structure of the data file and cause errors in processing the data. Because these responses have no analytic utility, they were identified and stripped from the SPSS data files before transfer to the operational folder.

There were also a number of variables that needed to be created or derived from existing variables which ETS uses to identify or track the data through the analytic processes. Because these variables have no intrinsic value outside of these processes, they were not provided to IEA for the master database but were only generated and retained in ETS operational data files. An SPSS macro was implemented to create and add these variables to the SPSS data files as they were transferred to the operational folder.

After the SPSS files were transferred to the operational folder, the last step in the process was to produce the ASCII extract data file and its accompanying XML data dictionary file.

22.3 Updating/adding data

The results of the several analytic processes at ETS produced new variables (or new data for existing variables) that required merging into the operational data files for internal quality and consistency checking before addition to the master database at IEA. These various data sources included, but were not limited to, the following activities (which are described elsewhere in the documentation):

- the production of scale scores for the literacy, numeracy and problem solving in technology-rich environments (PSTRE) components
- the development of indices for the skill use categories
- the derivation and imputation of income variables as specified by ROA
- the creation of variables to be used for trend analyses with the IALS and ALL surveys

Some of these data came in ASCII files that first needed to be converted to SPSS files before merging, some were already in SPSS file format, and some were represented in SPSS macro code that had to be applied to the operational SPSS files to be created and saved as separate files.

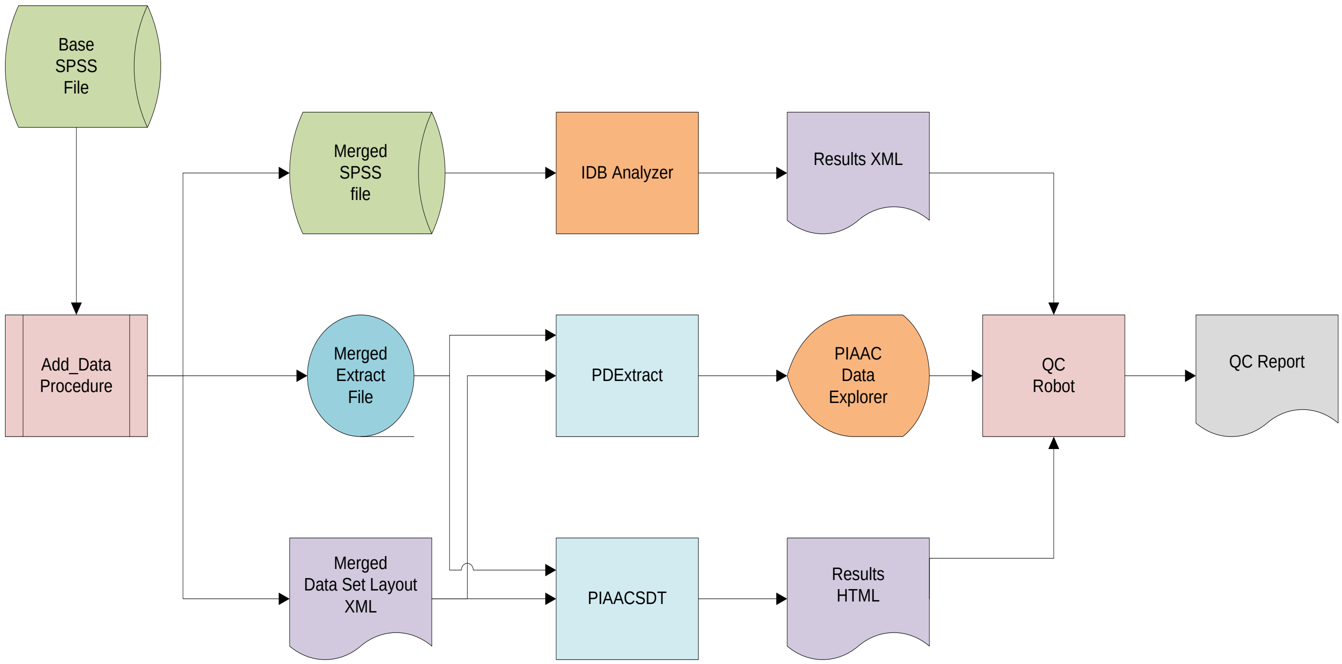
To efficiently, consistently and accurately perform these merging operations using a variety of input data sources, ETS developed a Python-based procedure that would iteratively process the data files for each country. Each application of the procedure only required as input a parameter file that specified the operations to be performed and the folder and file names for the input and output files.

The critical outputs for each application of this procedure were the SPSS file containing the new or updated variables, an SPSS file containing the merger of the operational file and the new or updated variables and the data and dictionary extracts of the merged SPSS file. Once these files were checked and approved by ETS, the SPSS file containing new or updated variables was sent to IEA for addition to or updating of the master database and the merge files became the new operational files.

22.4 Population and quality check of the PIAAC Data Explorer

The process to populate the PIAAC Data Explorer database and confirm the results it produces is summarized in Figure 22.1 below. For the purpose of explanation, consider that this process was applied separately to the data from each country.

Figure 22.1: PIAAC database population and quality control



The Base SPSS File contained the data as received from IEA/DPC and as forwarded to the appropriate country for its analysis and reporting.

The Add_Data procedure performed two functions. The first was conditional on whether a country provided supplemental data that was collected or derived and merged these data with the Base file. The second function created two files from the enhanced Base file: an ASCII text rectangular file containing the data values extracted from the Base file and an XML file containing information about the extracted data variables (location, format, labels). This Data Set Layout (DSL) XML is structured in a proprietary ETS schema.

The PDExtract program used the information from an input parameter file to process the data from the Extract file and metadata from the DSL file to produce a series of text files suitable for loading into the appropriate tables in the PIAAC Data Explorer (PDX) database. The program also produced a SQL script that is customized for performing the loading of these tables and contains a procedure for forming the data tables used by the PDX.

The PIAACSDT program also used the information from an input parameter file as well as a list of data variable names to calculate and produce summary data tables (SDT) – one analysis for each scale score. Each table in the analysis was a one-way tabulation of various statistics for each category of a given variable. The statistics pertained to a scale score and include percentage, average score and percentages within the benchmark levels. Each statistic was accompanied by the standard error estimate, degrees of freedom, number of cases on which the statistic is based and number of strata on which the standard error was based. All of these results were stored in an HTML document in full precision. This document may be viewed with any of the popular Internet browsers when accompanied by the appropriate Cascading Style Sheet (CSS) document, which ETS provides. The document may also be parsed or translated to produce Excel workbooks and report quality tables, among others.

In the QC Robot procedure, the results HTML document from the PIAACSDT program were used to generate analysis requests for the PDX, one for each variable, and the results returned from the PDX were compared with those in the HTML document. The results of these comparisons are posted to the QC Report document where differences above specified criteria are flagged and subsequently examined.

The only statistics that can be reported in the PDX which cannot be calculated by the PIAACSDT program are the percentiles. Because the calculation of the percentiles within the PDX uses more resources than the other statistics, only a subset of critical variables was selected for quality-assurance analysis. The IEA IDB Analyzer reads data from the Base SPSS file, uses SPSS macros to calculate the desired percentile statistics, and writes the results to an XML file. The QC Robot procedure processed this XML file in the same way as the HTML file from the PIAACSDT program and added the comparison results to the QC Report file.

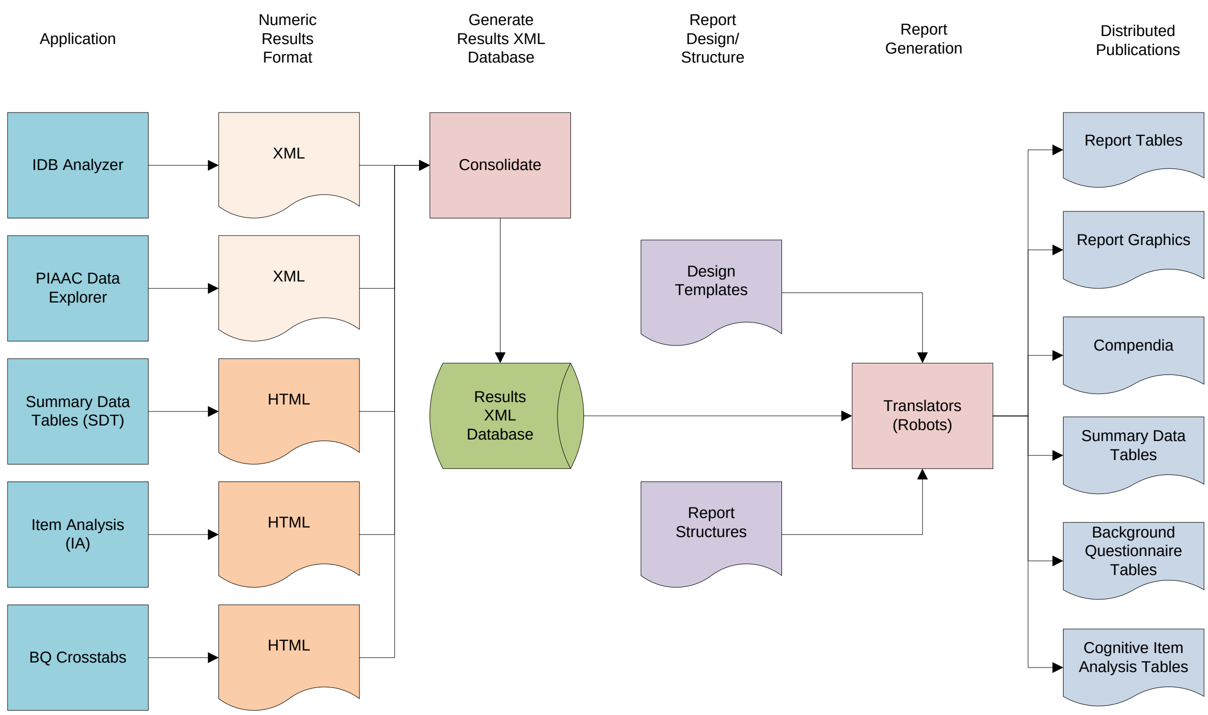
Prior to the first execution of the procedure described above, the IEA IDB Analyzer and the PIAACSDT programs were extensively calibrated with each other to ensure that the Merged SPSS and Merged Extract files were isomorphic and produced identical results for the statistics common to both programs.

22.5 Dynamic reporting system

The PIAAC dynamic report translation and publication system streamlined report and form generation by separating the data extraction and statistical computation process from the report design layouts and generation of publication formats. The generation process is shown in Figure 22.2. In the first stage, PIAAC-based Data Explorer, Data Analyzer, and procedural language applications performed data extraction and statistical computation across the entire PIAAC database and provided data files containing the numeric results for BQ and cognitive items in tagged language (XML, HTML) formats. In this stage, numeric computations were done once, and the numeric results were available to the second stage for efficient repurposing for quality control and report generation processing.

In the second stage, the publication system accepted report design and layout templates that were created with common desktop applications, and rendered XML structures based on those templates. Using these well-formed XML reporting structures, the system then applied XML-XSLT style sheet language or PYTHON-based scripts to transform the numeric results into viewable and publication-ready formats (PDF, RFT, Excel, HTML, and so on) for distribution. By increasing flexibility for rapid report generation customization through XML translation processing and the availability of a common numeric results archive, this two-stage phase approach to reporting dramatically reduced technical resource requirements and delivery times, enabling PIAAC to accommodate iterative data cleaning cycles while maintaining fixed publication delivery timelines.

Figure 22.2: PIAAC report translation and publication system



PIAAC BQ crosstabulations, summary data tables, item analyses, tables and graphical displays for possible inclusion in international and national reports, and compendia were generated by the Consortium using the PIAAC dynamic report translation and publication system. Following are descriptions and examples of each of these tables. All but the compendia are secure and not available for public view.

22.6 Summary data tables (SDT)

Via a secure FTP site, the Consortium delivered sets of files individually to each country containing summary data tables (SDT) that provided descriptive statistics for every categorical background variable in the respective country's PIAAC data file. For each country, the SDT included both international and idiosyncratic national background variables. The SDT were used by the Consortium and the countries for quality control and validation purposes: plausibility of 1) distributions of background characteristics and 2) performance results for groups, especially in the extent to which they agree with expectations or external/historical information.

For each variable, these tables contain weighted summary statistics, including variable identification, sample size, number of valid cases, weighted percentages of individuals corresponding to each valid response option, weighted percentages of individuals for whom none of the valid response options were selected, and within each categorical cell, the average score on one of the three PIAAC scale score domains. Standard errors were also included where applicable. An individual set of tables was provided for each scale score domain – literacy, numeracy, and for those countries that administered it, PSTRE. The SDT were provided in two formats – HTML and Excel.

The HTML files are suitable for viewing in a browser application, using the accompanying CSS that was provided. Two HTML files were provided for each of the three scales – literacy (LIT), numeracy (NUM), and problem solving (PSL) – separately by the set of international variables (INT) and the set of national adaptations and extension variables (NAT). The "INT" SDT files include the original BQ variables, the OECD-derived variables, and the quintile categorical variables derived from the skill use indices. The "NAT" SDT files include the original idiosyncratic national BQ variables. An additional analysis was performed for the reading component (RC) scores by selected BQ variables. When viewed in a browser, each file has a link at the top of the file to a Table of Contents at the bottom; after clicking on the link, a user can scroll to the left of the display to see links to each of the variables processed in the analysis.

Two types of Excel files were provided for each HTML file. These correspond to two modes of presenting the results:

1. A "Data" worksheet. Each row of the data sheet contains a statistic from the tables presented in the HTML file across all values within each variable and across all variables in the file. Each statistic is accompanied by its standard error estimate, estimated degrees of freedom, estimated population count (weighted N), and number of cases from the data file. The organization of this sheet allows for post-processing of the results by secondary analysis procedures.

2. "Report" worksheets. This consisted of one worksheet for each variable in the analysis, using the variable name as the name of the tab, and a sheet named Table of Contents that contains hyperlinks to the individual worksheets. Each variable-named worksheet contains the analysis results in tabular form, mimicking the tables in the HTML file display.

22.7 BQ crosstabulations

The BQ crosstabulations were produced for internal Consortium quality control and data validation during the initial stages of PIAAC data processing and cleaning. Their contents are similar to the SDT contents that were subsequently provided to the individual countries.

22.8 Item analysis tables, weighted and unweighted

Similar to the summary data tables, the item analysis tables contain summary information about the response types given by the respondents to the cognitive items. They contain, for each country, the percent of individuals choosing each option for multiple-choice items or the percent of individuals receiving each score in the scoring guide for the constructed-response items. They also contain the international average percentages for each response category. The item analysis tables were used by the Consortium and the countries for quality control, verifying data structure accuracy, and validation purposes. A brief description of the details of the calculation of item statistics for the PIAAC data follows.

PIAAC introduced many features for the first time to the large-scale population surveys of cognitive skills. Two main unique features that impact item analysis are: 1) the use of two modes of assessment – CBA and PBA, and 2) adaptive testing on computer.

Both features interact with the background characteristics and skills of the respondents who received particular sets of items. Even a simple statistic such as the proportion correct across two groups of respondents may not be directly comparable if, for example, it involves comparisons between groups taking two different modes of assessment, or groups following different adaptive-testing paths due to variation in skills. In general, younger and more educated respondents tended to receive CBA rather than PBA items based on their ICT skills. However, statistics for the items in a set administered to a group of respondents are comparable within a country. For example, item statistics of PBA items can be compared to each other. But because CBA items are clustered in smaller sets for multistage adaptive testing, direct comparison of item statistics among CBA items is limited.

All respondents with nonzero weights were included in the item analysis. Item analysis of cognitive data involves calculation of a set of statistics to describe the data in terms of quantity and quality before we apply any measurement model. Two sets of statistics were calculated on the unweighted data to represent the number of cases and structures of data using the uniform weight of 1, and also on the final weights to calculate similar statistics to describe the data in comparison to the reference of choice, such as international means.

Unweighted item analysis results are particularly useful to verify the accuracy of the data structure. Seven worksheets are provided in each Excel item analysis file for unweighted and weighted items: literacy core items, numeracy core items, literacy and numeracy PBA items, literacy CBA items, numeracy CBA items, PSTRE CBA items and reading component items.

Each worksheet has eight columns unless there are polytomous items in a set. Each row represents a unique item, identified in the first column entry with the item ID used throughout all phases of PIAAC. The second and third columns are the number of respondents for “not administered” and “not reached.” Due to the matrix sampling design, in addition to the two modes of administration, each respondent was given only a fraction of the items in the item pool. By design, these missing responses were termed “not administered.” In some cases, respondents were given tasks they did not attempt or reach during the time period allotted for the survey. Consecutively missing responses at the end of a block were termed “not reached.” Both “not administered” and “not reached” respondents are excluded in calculating percent correct.

In some cases, responses were missing because respondents chose not to perform a task. Any missing responses that were followed by a valid response (whether correct or incorrect) were termed as “omitted” responses for PBA items. This means a missing response on the last item in a PBA booklet was not treated as omitted. For the adaptively administered CBA items, the position of an item is not nearly as informative as the duration of time each respondent spent on it, as well as the type of input that the respondent provided using the keyboard or mouse. Clearly, the absence of keyboard or mouse responses from a respondent who skips items without having the chance to examine them is not a good indication of his or her skills. A heuristic decision was made that the absence of response when less than five seconds was spent on an item was treated as “not administered” even though it might have been followed by a valid response later on. Omitted responses were treated as wrong. The total consists of the sum of omitted, correct and incorrect responses. Percent correct is calculated as the number of respondents with the correct response divided by the total number of respondents who attempted the item.

Because statistically equivalent samples received either the literacy or numeracy PBA booklet, item statistics are comparable within a country, that is, an item with percent correct of 0.4 was more difficult than another item with percent correct of 0.65 for the PBA population. The comparability of PBA item statistics is limited across countries due to the population characteristics of the PBA respondents of each country, which is primarily driven by the ICT skills of respondents instead of good representation of the national population.

Using only the final weight to calculate item statistics means they are not comparable across countries due to the differential proportions of respondents who took a particular adaptive path. In particular, the total number correct would be greatly biased based on the distribution of paths. In order to increase comparability across countries, path weights were standardized using the international average of path proportions in addition to the final sample weights. The final weights (prior to the application of path proportions) were standardized to 5,000 for each country.

22.9 Compendia

Using the public-use files (PUF) as the source data, the compendia are sets of tables that provide categorical percentages for both cognitive and background items. The compendia are essentially redacted versions of the summary data tables. The purpose of the compendia is to support PUF users so they can gain knowledge of the contents of the PUF and use the compendia results to be sure that they are performing PUF analyses correctly. The item statistics reported in the compendia differ from the item analysis tables in two ways: 1) for confidentiality reasons, some countries have altered data or removed respondent records from their PUF files; and 2) the compendia do not use the routing methods employed in the item analysis. As a result, comparing compendia item statistics across countries for reporting purposes is not appropriate. The compendia reside on the OECD PUF Web site.

22.10 Report tables

The report tables are publication-ready tables that were provided by the Consortium to support the OECD international report. These tables were derived using the ETS Dynamic Reporting System. The data source is the PIAAC Data Explorer database. The PIAAC Data Explorer analysis and reporting engines generated the required reporting statistics.

Chapter 23: International Database and Data Analysis Tools

Ralph Carstens and Tim Daniel, IEA Data Processing and Research Center;
Eugenio Gonzalez, ETS

23.1 Overview

Designing, collecting, validating and analyzing PIAAC data was a very complex, highly demanding and collaborative process involving all Consortium partners, a broad range of external experts, all participating countries, and the OECD Secretariat. Naturally, this in turn led to a data product that reflects the design complexities. To support and promote secondary analyses, the OECD is making a public-use version of the international database and this technical report available to interested analysts and users in the scientific community as well as the general public. The international public-use version of the PIAAC database is made available in two different ways: i) as a database underlying a Web-based data analysis software, the PIAAC Data Explorer (PDX), and ii) a set of public-use files (PUF) which comprise person-level microdata from those countries that gave permission to release their national data.

This chapter is intended to provide a basic introduction to the PIAAC public-use database and the software tools capable of replicating the descriptive and inferential analysis presented in the initial publication “Skills Outlook 2013: First Results firm the Survey of Adult Skills (PIAAC)” (OECD, 2013). First, the chapter will discuss the contents of the public-use data both at the record as well as the variable level, the approach to identifying missing data under a complex, multi trajectory design, and the available database formats. Then, the chapter will describe general analytical considerations followed by the types of analysis supported by the two software tools provided by the Consortium: the International Data Explorer and the IDB Analyzer.

This chapter, however, does not intend to cover and illustrate the full range of possible analytical techniques appropriate for PIAAC and therefore does not describe, for example, advanced modeling of data such as structural equation modeling (SEM). Nonetheless, analysts wishing to use the public-use microdata to undertake advanced analysis not covered by the provided software or those wishing to use alternative statistical software packages will find sufficient technical information on the structure of the database, the included measures, and the variance estimation approaches to successfully configure such software and statistical models.

23.2 Files in the database

As described in Chapter 13 on data management, a large number of raw response data files and documentation were processed to form a series of files that jointly made up the national master databases for PIAAC, that is, all variables collected or derived as part of PIAAC. These national databases consisted of one main flat file holding respondent/household level information, a set of files holding information relating to the study of scoring reliability within and across countries, an audit log file holding interview process and timing data, and, for each respondent, a set of cognitive log files native to the CBA platform used in PIAAC. Of these files, only the main flat file is of key analytical interest and thus forms the basis of the public-use database described in this chapter. Other parts of the national master database, such as the cognitive log data, did not have a high analytical priority and in light of time and budgetary constraints are not part of the public-use data described here. However, the OECD may make derivatives of these files available to the public in the future.

At the time of processing, analysis, weighting, validation and reporting, all data for a particular PIAAC participant were kept separate from that of other participants. This partitioning per participant also holds for the PUFs and allows for a more flexible, staggered release of files to public users. This is especially useful given that a number of additional countries are currently implementing a second “round” of the first cycle of PIAAC. It is expected that these participants will be added to the public-use database in due time and be available through both the Data Explorer and in the form of a public-use microdata file. Further, certain PIAAC participants may require confidentiality agreements to be signed before public users may receive and use the data.[1] This and related information will be communicated by the OECD via the PIAAC website.

For the naming of physical files, lists of available samples and assigning value labels within the variables identifying countries and subnational entities, operational identifiers based on the ISO 3166/UN M49 standard were used. Table 23-1 provides details. Physical data files are named using the alpha-3 code of the national entity. Within databases, the variable CNTRYID holds the numerical codes and labels of the national entity to which the data belong. The variable CNTRID_E holds the numerical codes and labels of the subnational entity.

With the exception of three participants, it was a national entity that participated in the assessment; therefore, the codes and labels for CNTRYID and CNTRYID_E are identical. In the case of Belgium, only the Flemish part participated. In the case of Canada, the English- and French- speaking parts are identified as subnational entities. In the case of the United Kingdom, the database includes the data from two subnational entities: England and Northern Ireland. Keeping this information in two separate variables allows for analysis at the level of the national as well as subnational entities (as domains) as appropriate. The initial reporting in “Skills Outlook 2013: First Results firm the Survey of Adult Skills (PIAAC)” (OECD, 2013) was done at the level of national entities. Combined data for “England (UK)” and “Northern Ireland (UK)” was reported as “England/N. Ireland (UK)” in the international reporting. Data for Belgium (Flemish part only) was reported as “Flanders (Belgium).”

[1] At the time of writing, this applied to Australia.

Table 23-1: Operational participant codes and names used in PIAAC

National entity name	National entity numeric code	National entity alpha-3 code	Subnational entity name	Sub-national numeric code	Sub-national alpha-3 code
Australia	36	AUS	n/a	n/a	n/a
Austria	40	AUT	n/a	n/a	n/a
Belgium	56	BEL	Flanders (Belgium)	956	BFL
Canada	124	CAN	Canada (English)	1241	CEN
			Canada (French)	1242	CFR
Cyprus[2]	196	CYP	n/a	n/a	n/a
Czech Republic	203	CZE	n/a	n/a	n/a
Denmark	208	DNK	n/a	n/a	n/a
Estonia	233	EST	n/a	n/a	n/a
Finland	246	FIN	n/a	n/a	n/a
France	250	FRA	n/a	n/a	n/a
Germany	276	DEU	n/a	n/a	n/a
Ireland	372	IRL	n/a	n/a	n/a
Italy	380	ITA	n/a	n/a	n/a
Japan	392	JPN	n/a	n/a	n/a
Korea	410	KOR	n/a	n/a	n/a
Netherlands	528	NLD	n/a	n/a	n/a
Norway	578	NOR	n/a	n/a	n/a
Poland	616	POL	n/a	n/a	n/a
Russian Federation[3]	643	RUS	n/a	n/a	n/a
Slovak Republic	703	SVK	n/a	n/a	n/a
Spain	724	ESP	n/a	n/a	n/a
Sweden	752	SWE	n/a	n/a	n/a
United Kingdom	826	GBR	England (UK)	926	ENG
			Northern Ireland (UK)	928	NIR
United States	840	USA	n/a	n/a	n/a

23.3 Records in the database

This section describes the records included in the database. PIAAC used a highly complex assessment design that resulted in a number of possible trajectories through the interview process. It is therefore important for users to understand this design in order to make appropriate use of the database.

23.3.1 Records included in the database

As a general principle, each national master database and, by extension, each national public-use database includes the exact same records that were considered to be suitable for analysis. One

[2] Please refer to notes A and B regarding Cyprus in the Note to Readers section of this report.

[3] Please refer to the note regarding the Russian Federation in the Note to Readers section of this report.

exception to this rule in discussed below. More specifically, each record in the database generally corresponds to a responding sampled person. Each record in the database also conforms to the international target population definition, that is, adults between the ages of 16 and 65. All records in the database were adjudicated, weighted and used in the computation of response rates.

While the vast majority of records in the database are "true completes," that is, sampled respondents that followed the intended interview workflow until the end (regardless of administration mode and flow), there are noteworthy exceptions. The two main groups of respondents that are included in the database with weights and replicate but with only very little or partial information are i) literacy-related nonresponse cases, that is, respondents who were unable to take the assessment or discontinued it for one of three reasons,[4] and ii) certain types of break-offs, i.e., respondents who decided to discontinue the assessment after it commenced.

The inclusion of these two types of records directly relates to the PIAAC Technical Standard 4.3.3 (OECD, 2011b) that defines a "completed case." A completed case is one that minimally has:

a. responses to key background questions (age, gender, highest level of education and employment status) and a completed Core instrument (i.e., the interviewer asked the respondent all Core questions), or

b. responses to age and gender for literacy-related nonrespondents to the BQ and the Core instrument.

The original plan was to assign imputed scores at the lowest level of proficiency for these cases. However, this was not warranted from a psychometric point of view and the additional information reviewed. As a consequence, these types of records in the database are likely incomplete and not fully usable and, in the case of literacy-related nonresponse, will not have plausible values. In the analytical tools described below in this chapter, these cases will be reported as "not classified" in certain types of analysis.

23.3.2 Records excluded from the database

As part of the data collection, validation, and weighting, certain cases in the original master databases were excluded from the analysis and the public-use databases. The types of cases dropped from the databases include, but are not limited to: i) out of scope respondents, ii) households with no sampled persons, iii) noninterviews, meaning sampled persons who were not interviewed due to refusal or other reasons, iv) a small number of suspected falsified cases detected as part of the validation and quality control, v) respondents with less than the minimally required BQ items (age, gender, highest level of education and employment status) or age and gender in the case of literacy-related nonresponse, and vi) cases with certain anomalies or unclear origin. These cases were flagged accordingly and no weights were computed.

In relation to this, two notes should be made:

a. Two countries targeted respondents not part of the international target population definition (adults from 16 to 65). In the case of Denmark, this related to an oversample of Programme for International Student Assessment (PISA) students. In the case of Australia, an oversample targeted individuals at the age of 15 and between 66 and 75 years. Both groups of cases were excluded from the respective PUFs.

[4] The three types of literacy-related nonresponse are: i) language problems (disposition code 7), ii) reading and writing difficulty (code 8), and iii) learning/mental disability (code 9).

b. In the case of Canada, disclosure risk assessment demanded the reweighting of a small number of cases from a particular domain of respondents in order to comply with Statistics Canada's minimum weight reporting standards. As a consequence, some cases were excluded from the public-use microdata file for Canada and its corresponding weights were loaded onto other cases in the domain. This means the full set of cases used for the international reporting and the revised set of cases included in the PUF for Canada are not identical. Therefore, it will be impossible to replicate reported estimates precisely using the PUF. However, these small weight adjustments should have no practical relevance and should not affect the agreement of estimates published by the OECD, those produced by the Data Explorer, and those made on the basis of the PUF.

23.4 Variables in the database

The PIAAC design is a highly complex one that integrates sophisticated sampling and weighting approaches, a multitrajectory assessment, rich BQ, CBA and PBA modes, innovative item formats, related process information, and a range of derived measures, indicators and indices. In total, each national database includes 1,712 common variables. There were a total of 1,206 country specific variables.

With that said, it is obvious that such a rich database contains variables of varying analytical utility and priority. For example, a large number of variables only include process-related information or temporary information that is necessary, for example, during the computation of weights. This section therefore describes the types of variables included in each participant's public-use database. It also describes those excluded because they carry no analytical utility for international comparisons or address identified and/or assumed disclosure risks.

23.4.1 Variables included

The public-use database underlying the PDX and the PUF contains different sets of variables. The PUF includes a comprehensive set of 1,328 variables. Of these, only 575 are included in the Data Explorer database, implying that certain sets are not informative for analysis in the PDX yet are included in the PUF for secondary analysis. The majority of variables included only in the PUF relate to the individual cognitive item scores and process information. Table 23-2 provides a breakdown of variables by type, name or naming convention, and whether the respective group is available in the PDX or the PUF.

Table 23-2: Variable groups and their description, count, naming convention, and inclusion in public-use database

Variable group	**Description**	**Count**	**Names or naming convention**[5]	**Inclusion**
Identifiers	National entity, subnational entity and respondent identifier	3	CNTRYID, CNTRYID_E, SEQID	DX and PUF
Resolved demographics	Resolved age and gender	2	AGE_R, GENDER_R	DX and PUF
Derived disposition codes	Summary disposition codes derived from detailed disposition codes	3	DISP_CIBQ, DISP_MAIN, DISP_MAINWRC	DX and PUF
BQ	Originally collected BQ responses (after mapping from national data where applicable)	249	{A-J}_{Q/D}*{a-m}*, e.g., B_Q01a	DX and PUF
BQ – Coded responses	Coded values for respondents' language, education, occupation, industry, country, and region	13	LNG_*, ISCED_HF, ISCO08_*, ISIC4_*, CNT_*, REG_TL2	DX and PUF
BQ – Derived background information	Background information derived from original or coded BQ items	30	AGE10LFS, AGEG5LFS, BIRTHRGN, BORNLANG, CTRYQUAL, CTRYRGN, FIRLGRGN, FORBILANG, FORBORNLANG, HOMLANG, HOMLGRGN, IMGEN, IMPAR, IMYRCAT, IMYRS, ISCO*, ISCOSKIL4, ISIC*, NATBILANG, NATIVELANG, NOPAIDWORKEVER, PAIDWORK12, PAIDWORK5, SECLGRGN,	DX and PUF
BQ – Derived education information	Education information derived from original or coded BQ items	26	AETPOP, EDCAT*, EDWORK, FAET*, FE12, FNFAET*, FNFE12JR, LEAVEDU LEAVER1624, NEET, NFE*, PARED, YRSQUAL, YRSGET, VET	DX and PUF
BQ – Derived earnings information	Earnings variables (continuous, continuous purchasing power parity (PPP) corrected, deciles) for BQ earnings items	17	EARN*, MONTHLYINCPR, YEARLYINCPR	DX and PUF
BQ – Derived skill use information / scale scores	Scales scores (standardized and categorized weighted likelihood estimation) for skill use items in BQ	26	LEARNATWORK*, READYTOLEARN*, ICTHOME*, ICTWORK*, INFLUENCE*, NUMHOME*, NUMWORK*, PLANNING*, READHOME*, READWORK*, TASKDISC*, WRITHOME*, WRITWORK*	DX and PUF
BQ – Derived trend information	Recoded versions of BQ responses to facilitate trend analysis with IALS/ALL data	44	As for original BQ variables yet with suffix "_T" or "T1"	DX and PUF

[5] {Brackets} indicate the possible characters used in variable names. Asterisks (*) indicate name stems.

Variable group	Description	Count	Names or naming convention[5]	Inclusion
BQ – Derived coarsened information	Coarsened versions of BQ responses (collapsed, categorized or top-coded)	29	As for original BQ variables yet with suffix "_C"	DX and PUF
BQ – Derived cognitive routing	Variables derived from BQ at the time of collection to determine adaptive routing	3	COMPUTEREXPERIENCE, NATIVESPEAKER, EDLEVEL3	PUF only
Cognitive scores, pass flags, random numbers	Core scores, pass status, and random module allocation recorded at the time of collection	13	CBA_CORE_STAGE*_SCORE, CORESTAGE*_PASS, RANDOM_CBA_*, CBA_START, PPC_SCORE, RANDOM_PP	PUF only
Cognitive routing – Derived	Variables derived from the actual routing describing the module allocation	9	PAPER, CBAMOD*, PBROUTE	DX and PUF
Observation module	Interviewer's descriptions of the assessment session	13	ZZ*	PUF only
Cognitive item responses and process information	Cognitive item information: actual response (R), scored response (S), total time (T), time to first action (F), number of actions (A)	720	{C/D/E/M/N/P/U}*{A/F/R/S/T}, e.g., C301C05S	PUF only
Numeracy, literacy and problem-solving scale score status	Status flags indicating availability of scale scores for the respective domain	3	LITSTATUS, NUMSTATUS, PSLSTATUS	DX and PUF
Numeracy, literacy and problem-solving scale scores	Scale scores (plausible values) for each of three domains	30	PVLIT1 to PVLIT10, PVNUM1 to PVNUM10, PVPSL1 to PVPSL10	DX and PUF
Reading components scores	Total correct scores (point estimates) for reading components	3	PRC_PV_SCR, PRC_SP_SCR, PRC_PC_SCR	DX and PUF
Reading components timers	Timing values for reading component parts	5	PRC_PV_Q1, PRC_SP_Q1, PRC_PF_Q1, PRC_PF_Q2, PRC_PF_Q3	DX and PUF
Variance estimation	Variables controlling variance estimation stratification, method, and number of replicates	6	VEMETHOD, VEMETHODN, VEFAYFAC, VENREPS, VARSTRAT, VARUNIT	DX and PUF
Full weight and replicates	Complex sample estimation weights	81	SPFWT0, SPFWT1 to SPFWT80	DX and PUF
Total		1,328		

23.4.2 Variables excluded, suppressed or coarsened for some or all countries

The public-use databases only include a subset of the information available in the master databases. The public-use database does not include any data collected using national adaptations and extensions. It only includes data that were collected or derived across all countries. Further, a sizable number of variables were excluded in consultation with the OECD Secretariat and the BPC because they i) have no or little analytical utility, ii) were intended for internal or interim purposes only, iii) relate to secure item material, or iv) include personally identifiable data, or at least data that may increase the risk of unintended or indirect disclosure.

The groups of variables excluded from the public-use database are:

a. direct, indirect, and operational identifiers for respondents, interviewers, scorers, key operators, and paper materials

b. interim sampling, disposition, data availability, demographic, and weighting information

c. certain BQ or process variables that are available in coded or derived form (for example, country and language), especially detailed write-ins

d. all national adaptations and extensions in the BQ

e. interviewer's scoring of paper-based core items

f. detailed response information for secure problem-solving items

g. original scale score values (theta) before standardization to an international metric

National data is not of key interest in an international large-scale assessment and comparison. However, national data might be available by directly contacting the concerned PIAAC participant.

A particularly important issue is to preserve the confidentiality of individual respondents in the release of the public-use aggregate (PDX) and microdata (PUF) in order to prevent unintended or indirect disclosure. The risk of such disclosure is greatest in cases where the combined characteristics of a respondent in a sample lead to a unique individual in the population. The higher the sampling fraction, the more likely a unique record in the sample will also be unique in the population. As agreed by the BPC, countries were given the possibility to either coarsen or suppress their data prior to submission to the Consortium and the OECD and/or afterward during the production of the public-use database. PIAAC participants were asked to suppress information only when deemed absolutely necessary to meet national legislative requirements.

The database underlying the PDX and PUF was subject to around 700 instances of suppression (participant x variable) at the cell or column level. The majority of these instances relates, but are not limited, to:

a. detailed age

b. detailed language, country of birth, or region information

c. detailed education information (BQ section B)

d. detailed occupation (ISCO) and industry (ISIC) information

e. detailed, original, or derived earnings variables (BQ section D)

f. variance strata and unit information

Suppressed data are represented in the database by means of missing codes. As with national data, more detailed data might be available directly from the concerned participant.

Database users should note that the most complete set of information was available to the Consortium for analysis and the OECD for reporting and archiving. The PDX is based on a reduced database, that is, it includes fewer variables and less information as a function of suppressions. Finally, the PUF is the most restricted database in PIAAC.

In almost all cases where more than one participant requested the suppression of a particular variable for the PDX or PUF, a coarsened version of this variable (suffix "_C") was created that includes the level of detail deemed suitable for public release by the concerned countries (see group "BQ – Derived coarsened information" in Table 23-2 above). Analysts are therefore recommended to use such a coarsened variable if the aim of the analysis is to include the most complete set of countries, albeit with a reduced level of detail.

As a result (and similar to other data collections), public users of the databases in the PDX or PUF may be unable to fully replicate particular tables, figures, and other exhibits in the international reporting because such reporting was based on the most complete set of confidential information, which is not available to the general public.

23.5 Representing valid and missing data

As in all survey projects, missing data is a natural phenomenon. PIAAC is no exception, and despite the intention to collect complete or almost complete information, there are related gaps in the database. In principle, missing data in a survey may occur when there are no or almost no observed data as well as no administrative data for a respondent (unit nonresponse) or when some variables for a respondent are unknown or cannot be known (item nonresponse). Missing data can further be distinguished semantically in two broad groups: i) data that cannot exist due to the way a survey is designed and ii) data that were supposed to be observed but were not.

To understand the missing data pattern in PIAAC, users are reminded of the complex assessment design. Missing data in PIAAC can occur for a number of reasons. The main ones are:

a. Data are missing by design (that is, it is known a priori they will not collected) for some or all respondents because of the way the assessment is designed.
    i. Respondents with literacy-related dispositions (see above) were not administered the interview.
    ii. A small number of PIAAC participants did not participate in one or both of the international options: i) problem solving in technology-rich environments and ii) reading components.
    iii. Certain sections in the BQ were intentionally presented to subpopulations (domains) only with reference to responses given to prior questions (“valid skip”).
    iv. Respondents were by default administered the CBA or, as a result of their lack of computer familiarity, inability or refusal to take the exercise on the computer and/or performance on core modules, a full or reduced PBA was administered.
    v. Respondents following the paper-based path were not administered problem-solving items and therefore have no plausible values for problem solving.
    vi. Domain item clusters (CBA and PBA) were assigned based on random allocation and previous proficiency information collected (in the case of CBA).
b. Data are missing as a result of the response process.
    vii. Respondents may have broken off the interview after it was started as a function of, for example, time, motivation, fatigue, or sensitive questions being asked.
    viii. Respondents may have explicitly refused (“refused”) to respond to questions in the BQ or they may not have known the answer to a question with sufficient certainty (“don’t know”).
c. Data in a few instances are missing due to logistics, processing, or analysis.

ix. Data were captured yet paper booklets and/or CBA result files were lost during transfer.

x. Erroneous routing in national versions of the BQ collected fewer data items for particular respondents than intended.

xi. Certain data items (variables and/or a subset of values) were not provided or suppressed due to regulations relating to confidentiality of information.

xii. Respondents with literacy-related dispositions (see above) were usually not assigned domain scores.

xiii. A small number of values were obvious outliers, otherwise useless, or erroneously coded in the original national databases.

It should be noted that no imputation was intended for missing item responses except for i) the imputation of earnings from precise and/or broad categories, and ii) the multiple imputation of proficiency scale scores for the literacy, numeracy and problem-solving domains.

Table 23-3 below provides an overview of the main missing values and their semantic, scope and representation in SAS and SPSS PUFs. The representation of missing values differs in these two statistical packages. In SAS, the standard missing code (.) and special missing values (.A thru .Z) were used. In SPSS, a “dynamic” code that depends on the length of the numeric variable was used. Variables of length 1 use missing values 6, 7, 8 and 9; those of length 2 use missing values 96, 97, 98 and 99; variables of length 3 use 996, 997, 998, 999; and so on unless missing values conflicted with payload values, in which case the variable lengths were increased.

The PIAAC public-use databases also include a small number of coded variables that are defined as strings because the respective coding schemes are defined as string. For example, occupational codes may appear as using a numerical scheme but need to be stored as strings because codes include leading zeros that would be lost if converted to a number. The use of string variables and, therefore, string missing values relates to: i) ISCO codes for occupation, ii) ISIC codes for industry, iii) region codes, and iv) language codes. In these cases, number-based strings such as “9999” were used to represent missing data.

Table 23-3: Generally used missing values in the public-use database (DX and PUF)

Semantic	Scope	Label	SAS	SPSS
Valid skip	BQ and any variables derived from it; reading components	"Valid skip"	Numeric: .V String: "996," "9996"	Numeric: 6, 96 … String: "996," "9996"
Don't know	BQ and variable derived from it	"Don't know"	Numeric: .D String: "997," "9997"	Numeric: 7, 97 … String: "997," "9997"
Refused	BQ and variable derived from it	"Refused"	Numeric: .R String: "998," "9998"	Numeric: 9, 98 … String: "998," "9998"
Not stated/inferred, invalid, not codeable, omitted, not provided, or suppressed	Almost all variables	"Not stated or inferred" (general) "Not reached/Not attempted" (cognitive items)	Numeric: .N String: "999," "9999," "99999"	Numeric: 9, 99 … String: "999," "9999," "99999"
Not administered / not applicable (missing by design)	Cognitive items	n/a	Numeric: (.)	Numeric: (.)

In addition to the general missing scheme described above, which applies to the largest set of variables, the specifications of some derived variables included missing schemes specific to a particular variable or, in some cases, a small set of variables. These missing values are fully documented in the SPSS files and SAS format scripts. Given that the number space for missing values (or letters in case of SAS special missing values) is limited, some of the per-variable missing schemes may use the same missing code, yet the semantic of these codes may vary from one variable to the next. Database users are strongly encouraged to review the coding of missing values in derived BQ variables carefully, using the information provided as part of the SAS/SPSS files and earlier in this report prior to analysis.

23.6 Public-use file (PUF) formats

While the database underlying the Data Explorer is not directly accessible to users, the PUFs are. They are being made available in two standard formats – SPSS and SAS – allowing for data to be loaded and used in these and many other standard packages.

SPSS data files are standard, Windows-based .sav files and encoded in Unicode (UTF-8). SPSS data files include full dictionary information from the applicable metadata maintained in the codebooks: i) variable types and formats, ii) variable labels, iii) value labels (including any missing value labels), iv) missing value definitions, and v) variable measurement levels.

SAS formatted files are standard, compressed .sas7bdat data files for Windows environments and encoded in Unicode (UTF-8). Variable types, widths, decimals, and labels are assigned to all variables according to the labels defined in the metadata. SAS does not provide for a way to permanently stored value labels on the file. Therefore, each PUF file in SAS format is accompanied by an equivalently named .sas file that includes syntax to assign formats (value labels). The SAS format syntax files include the relevant LIBNAME (in), PROC FORMATS, DATA and FORMATS statements. These syntax files can be executed against each individual SAS export file in order to display value labels in analytical procedures such as PROC UNIVARIATE, PROC FREQ, and so on.

23.7 Data analysis and software tools

23.7.1 General considerations for data analysis using PIAAC data

For analysts familiar with population estimation using other large-scale educational survey databases such as those produced by, for example, the OECD PISA program or IEA studies, the analysis of PIAAC data will present relatively few difficulties after becoming familiar with the conceptual foundation and the methodological, operational, and analytical details of the study, especially the BQ framework (OECD, 2011a) and the BQ itself (OECD, 2010). For those unaccustomed to working with complex survey sample data, the technical report as a whole, this chapter in particular, and the analytical tools provided by the Consortium should contain sufficient technical information and references to support statistically correct analysis.

The three main analytical requirements that any analysis of PIAAC data needs to account for are i) the use of sampling weights, ii) the complex multistage cluster sample design that was implemented to balance the research goals and cost-efficient operations, and iii) the use of multiply imputed proficiency estimates, the so-called "plausible values." The key challenge for analyzing PIAAC data, especially when one or more of the proficiency scales are involved, lies at the intersection of the uncertainty in estimating population characteristics due to sampling and the uncertainty introduced by the use of multiple imputations. In addition, another key challenge for PIAAC – in contrast to other international studies – is that there was not a common variance estimation procedure across all participating countries. Chapters 14 and 15 include details of the sampling, weighting and variance estimation techniques intended for PIAAC, the approach adopted by each country, and the mathematical combination of sampling and imputation variance. Chapter 17 includes details on the IRT and latent regression models used in deriving plausible values.

Standard analytical packages for the social sciences and educational research do not readily recognize or support handling the complex sample and assessment design. This gap is filled by the two software tools made available by the Consortium to assist database users to access and analyze PIAAC data and produce basic outputs: i) the PIAAC Data Explorer (PDX) and ii) the IEA's International Database Analyzer (IDB Analyzer). Each of these two software tools addresses a slightly different set of needs. While the PDX is a web based application that allows relatively easy and publication-ready access to basic estimates of means, totals and proportions, the IDB Analyzer used in conjunction with the PUFs provides unit record access to the public-use database and the opportunity to conduct analysis offline, derive additional variables, and produce various estimates for further use and reporting. The PDX and the IDB Analyzer are described in turn in the remainder of this chapter.

A variety of statistical software packages are available as alternatives for the analysis of complex sample data with support for the jackknife and/or BRR replication methods implemented in PIAAC. Still, all of these packages would require participant-by-participant runs or custom scripting to configure the variance estimation used by each participant. Further, these packages may or may not support the simultaneous integration of sampling and imputation variance as required when using PIAAC data. WesVar (2008; Westat Inc., 2007, 2008) software for complex sample analysis is available free of charge from Westat. Commercial packages that include support for the weighting and replication methods used in PIAAC, among others, are SAS 9.4, SUDAAN 11 (2013), and Stata 13 (2013).[6] A (slightly outdated) feature comparison of these packages is available in Heeringa, West and Berglund (2010, Appendix A, pp. 399+).

[6] Only current versions are mentioned. Previous versions may have support for complex sample analysis.

More generally, a detailed description of contemporary sampling and weighting approaches as well as analytical approaches and techniques for complex survey data analysis can be found in Heeringa, West, and Berglund (2010). Additional analytical advice and background useful for PIAAC and the previous international adult assessments may also be found in the user guides for IALS (Statistics Canada, 2001) and ALL (Statistics Canada, 2002) studies.

23.7.2. ETS PIAAC Data Explorer (PDX)

The PDX is a web based application developed by ETS that allows the user to query the PIAAC International Database via a web browser. The PDX can be used to compute a diverse range of statistics including, but not limited to, means, standard deviations, percentages by subgroup, percentages by levels, and percentiles. All statistics are computed taking into account the sampling and assessment design. In addition, the PDX has the capability of conducting significance testing between statistics from different groups, and displaying the results in graphical form. Results from the PDX can be directly exported and saved is MS Word, MS Excel and HTML formats. The PDX is accessible from any computer connected to the internet from the following address: http://piaacdataexplorer.oecd.org/ide/idepiaac.

23.7.3 IEA IDB Analyzer

The IEA International Database Analyzer (IDB Analyzer, 2013) is an application developed by the IEA Data Processing and Research Center (IEA DPC) in Hamburg, Germany. The IDB Analyzer can be used to combine and analyze data from IEA's large-scale assessments such as Trends in International Mathematics and Science Study (TIMSS), TIMSS Advanced, Progress in International Reading Literacy Study (PIRLS), the Second Information Technology in Education Study (SITES), the Teacher Education and Development Study in Mathematics (TEDS-M), the Civic Education Study (CivEd), and the International Civic and Citizenship Education Survey (ICCS) as well as analyze data from the OECD's Teaching and Learning International Survey (TALIS), PISA and PIAAC.

The IDB Analyzer creates SPSS syntax that can be used to perform analysis with these international databases. In other words, it requires SPSS (Version 15 or above) to be installed on the user's system. The syntax generated and the referenced macros take into account information from the sampling design in the computation of sampling variance. In addition, it handles plausible values. The resulting code can be used to calculate estimates of achievement and their corresponding standard errors, combining sampling and imputation variance. The code generated by the IDB Analyzer enables the user to compute descriptive statistics and conduct statistical hypothesis testing among groups in the population without having to write any programming code.

The IDB Analyzer is licensed free of cost by the IEA and is for use only in accordance with the terms of the licensing agreement. While users can use the software for free, they do not have any ownership of, copyright or other intellectual property rights to the software itself or its components, including the SPSS macros. Users are only licensed to use the SPSS enclosed macros in combination with the IDB Analyzer unless explicitly authorized by the IEA in writing.

The Analyzer is available from the following permanent URL: http://www.iea.nl/data.html. The software license expires at the end of each calendar year, when users will again have to download and reinstall the most current version of the software. Features will be added on a continuous basis to support additional surveys and databases or include additional types of analysis, options or outputs. Technical support for the IDB Analyzer can be obtained by contacting the IEA Data Processing and Research Center's Software Unit at software@iea-dpc.de.

The IDB Analyzer is fully self-documenting, and each version comes with a comprehensive help manual as part of the installation (International Association for the Evaluation of Educational Achievement [IEA], 2013). Users of the PUFs (PUF) are referred to this more detailed documentation with respect to the use and interpretation of the Analyzer's features, options, and outputs.

The IDB Analyzer consists of two modules – the Merge Module and the Analysis Module – which are integrated in one common application window.

23.7.3.1 The Merge Module

The Merge Module is used to combine data files from different study participants, and when necessary, merge data files from different sources like student BQs and achievement files, or student background files with teacher- or school-level files. The Merge Module is only available to use with IEA databases and others in which the data are published separate by participant, currently TALIS and PIAAC. In the case of PIAAC, each participant corresponds to a single, flat data file and hence only requires vertical merging, that is, that of one or more participants.

The Merge Module also allows the user to easily select individual or groups of variables to create a smaller and more manageable dataset. When running the Merge Module, the IDB Analyzer creates SPSS code that merges and combines files specified by the user, keeping only the selected variables yet automatically adding all mandatory variables for correct variance estimation.

Merged data files created using the Merge Module can be processed either with the Analysis Module (see below) of the IDB Analyzer or by any other analysis software that accepts SPSS files as input.

23.7.3.2 The Analysis Module

The Analysis Module of the IDB Analyzer provides procedures for the computation of means, percentages, standard deviations, correlations, and regression coefficients for any variable of interest overall for a participant, and for specific subgroups within a participant. It also computes percentages of respondents in the population that are within, at, or above benchmarks of performance or within user-defined cut points in the proficiency distribution, percentiles based on the achievement scale, or any other continuous variable.

The Analysis Module can be used to analyze data files from the above mentioned studies, regardless of whether they have been preprocessed with the IDB Analyzer Merge Module. The Analysis Module can create code for several analysis procedures. Like the Merge Module, the Analysis Module creates SPSS code that computes the statistics specified by the user.

The following analyses can be performed with the Analysis Module:

a. Percentages and means: Computes percentages, means and standard deviations for selected variables by subgroups defined by the user. The percentage of missing responses is included in the output.

b. Percentages only: Computes percentages by subgroups defined by the user.

c. Regression: Computes regression coefficients for selected variables predicting a dependent variable by subgroups defined by the user. New in this version of the IDB Analyzer is the capability of including plausible values as dependent or independent variables in the regression equation. When more than one set of plausible values is specified in the analysis, the analysis is carried out using the first of the plausible values, then the second, and so on. In the end, the results are summarized across the plausible values.

d. Benchmarks: Computes the percentage of students meeting a set of user-specified performance or achievement benchmark by subgroups defined by the user. It computes these percentages in two modes: cumulative (percentage of students at or above given points in the distribution) or discrete (percentage of students within given points of the distribution). It can also compute the mean of an analysis variable for those at a particular achievement level. As an additional feature, the IDB Analyzer allows the user to compute percentages of people at each of the proficiency levels including, or excluding those that did not participate in the assessment.

e. Correlations: Computes correlation for selected variables by subgroups defined by the grouping variable(s). New in this version of the IDB Analyzer is the capability of computing the correlation between plausible values. When more than one set of plausible values is specified in the analysis, the analysis is carried out using the first of the plausible values, then the second, and so on. In the end, the results are summarized across the plausible values.

f. Percentiles: Computes the score points that separate a given proportion of the distribution of scores by subgroups defined by the grouping variable(s).

Prior to every analysis, the Analyzer calculates unweighted and weighted descriptive statistics for the analysis variables (means, standard deviations, minimum and maximum), and frequencies by analysis subgroups. In addition, except when computing percentiles, the estimate of the population size for each of the subgroups processed (sum of the sampling weights) and the corresponding standard errors are computed. Bar or line charts are drawn by default when computing percentages, percentages and means, and when calculating the percentages of the population within benchmarks with or without an analysis variable.

When calculating these statistics, the IDB Analyzer has the capability of using any continuous or categorical variable in the database, or makes use of scores in the form of plausible values. When using plausible values, the IDB Analyzer generates code that takes into account the multiple imputation methodology in the calculation of the variance for statistics as it applies to the corresponding study.

All procedures offered within the analysis module of the IDB Analyzer make use of appropriate sampling weights and standard errors of the statistics that are computed according to the variance estimation procedure required by the design as it applies to the corresponding study. In the case of PIAAC, this functionality extends to the level of participants, as the variance estimate method (VEMETHOD) and number of replicate weights (VENREPS) is encoded in the respective PUF.

For a complete list of features, options, and output fields and parameters, users are referred to the help manual that is part of every installation.

References

Heeringa, S. G., West, B. T., & Berglund, P. A. (2010). *Applied survey data analysis*. Boca Raton, FL: Chapman & Hall/CRC. See also http://www.isr.umich.edu/src/smp/asda/

IDB Analyzer (Version 3.1) [Computer software], (2013). Retrieved from the International Association for the Evaluation of Educational Achievement website: http://www.iea.nl/data.html

International Association for the Evaluation of Educational Achievement. (2013). *Help manual for the IDB Analyzer* (Version 3.1). Hamburg, Germany: IEA Data Processing and Research Center.

Organisation for Economic Co-operation and Development. (2010). *PIAAC background questionnaire* (MS Version 2.1). Retrieved from the OECD website: http://www.oecd.org/edu/48442549.pdf

Organisation for Economic Co-operation and Development. (2011a). *PIAAC conceptual framework of the background questionnaire. Main survey.* Retrieved from the OECD website: http://www.oecd.org/site/piaac/PIAAC(2011_11)MS_BQ_ConceptualFramework_1%20Dec%202011.pdf

Organisation for Economic Co-operation and Development. (2011b). *PIAAC technical standards and guidelines. December 2011*. Paris: Author.

Organisation for Economic Co-operation and Development. (2013). *Skills Outlook 2013: First Results firm the Survey of Adult Skills (PIAAC).* Paris, France: OECD.

Stata 13 [Computer software], (2013). College Station, TX: StataCorp LP.

Statistics Canada. (2001). *International Adult Literacy Survey. Microdata user's guide*. Ottawa, Canada: Statistics Canada.

Statistics Canada. (2002). *The Adult Literacy and Life Skills Survey, 2003. Public use microdata file user's manual*. Ottawa, Canada: Statistics Canada.

SUDAAN 11 [Computer software], (2013). Research Triangle Park, NC: RTI International.

Westat Inc. (2007). *WesVar 4.3 user's guide*. Rockville, MD: Author.

Westat Inc. (2008). *Addendum to the WesVar user's guide. New features in WesVar 5.1*. Rockville, MD: Author.

WesVar (Version 5.1), Replication-Based Variance Estimation for Analysis of Complex Survey Data [Computer software], (2008). Rockville, MD: Westat Inc.

Appendix 1: PIAAC Main Study Item Pool Characteristics: Literacy, Numeracy and PSTRE

PIAAC PSTRE Items

Item Difficulty (RP67)	Unit ID and Name	Proficiency Level (RP67)	Other ID	Slope	Difficulty	t1	t2	t3
374	U04A - Class Attendance	3	U04A000P	0.426	1.666	-6.277	3.573	2.704
355	U11B - Locate Email - File 3 emails	3	U11B000P	0.471	0.774	-5.357	4.062	1.295
346	U02 - Meeting Rooms	3	U02X000P	1.184	0.784	-0.271	0.500	-0.229
342	U06A - Sprained Ankle - Site Evaluation Table	3	U06A000S	1.132	1.000			
325	U06B - Sprained Ankle - Reliable/Trustworthy Site	2	U06B000S	0.534	-0.244			
321	U23 - Lamp Return	2	U23X000P	0.533	-0.052	-5.338	3.885	1.452
320	U21 - Tickets	2	U21X000S	1.191	0.310			
316	U03A - CD Tally	2	U03A000S	1.274	0.223			
305	U07 - Digital Photography Book Purchase	2	U07X000S	1.104	-0.237			
299	U01B - Party Invitations - Accommodations	2	U01B000S	1.531	-0.286			
296	U19B - Club Membership - Eligibility for Club President	2	U19B000P	1.072	-0.677	-0.387	0.387	
286	U16 - Reply All	1	U16X000S	1.377	-0.773			
286	U01A - Party Invitations - Can/Cannot Come	1	U01A000P	0.755	-1.047	-1.933	0.987	0.945
268	U19A - Club Membership - Member ID	1	U19A000S	1.414	-1.367			

PIAAC Numeracy Items

Item Difficulty (RP67)	Unit ID and Name	Trend Status	PIAAC Item ID	Proficiency Level (RP67)	Other ID	Slope	Difficulty
375	612 - Dioxin (MOD)	New	C612A518	4	C612518	0.981	1.930
354	632 - EducationalLevel	New	C632P001	4	E632001	0.632	1.313
348	610 - CompoundInterest	Trend	P610A515	4	M610515	1.698	1.619
341	623 - Wine	Trend	P623A618	4	M623618	1.414	1.452
332	660 - Weighthistory	New	C660P004	4	E660004	0.805	1.059
326	665 - Coopertest	New	C665P002	4	E665002	1.255	1.129
324	641 - Amoeba	New	C641P001	3	E641001	1.167	1.081
320	624 - BMI	Trend	C624A620	3	C624620	1.398	1.057
318	634 - Peanuts	New	C634P002	3	E634002	1.639	1.064
317	644 - NZExports	New	C644P002	3	E644002	1.650	1.051
315	661 - Studyfees	New	C661P002	3	E661002	1.155	0.916
315	657 - Package	New	C657P001	3	E657001	0.642	0.626
314	651 - Fertilizer	New	C651P002	3	E651002	1.509	0.973
308	661 - Studyfees	New	C661P001	3	E661001	1.392	0.847
308	620 - Inflation	Trend	C620A612	3	C620612	0.878	0.660
307	664 – Orchestra Tickets	New	C664P001	3	E664001	1.333	0.819
305	634 - Peanuts	New	C634P001	3	E634001	1.150	0.719
303	617 - Map	Trend	C617A605	3	C617605	1.067	0.653
301	622 - Classified	Trend	C622A615	3	C622615	0.851	0.533
297	618 - SixPack1	Trend	C618A608	3	C618608	1.024	0.543
296	611 - TempScale	Trend	C611A517	3	C611517	0.847	0.439
294	636 - LabReport	New	C636P001	3	E636001	0.870	0.405
287	617 - Map	Trend	C617A606	3	C617606	0.794	0.233
282	619 - Tiles	Trend	C619A609	3	C619609	1.087	0.279
276	623 - Wine	Trend	C623A617	3	C623617	1.327	0.238
276	660 - Weighthistory	New	C660P003	3	E660003	0.936	0.105
273	606 - Solution	Trend	C606A509	2	C606509	1.051	0.107
267	620 - Inflation	Trend	C620A610	2	C620610	1.365	0.097
266	632 - EducationalLevel	New	C632P002	2	E632002	0.938	-0.071
261	611 - TempScale	Trend	C611A516	2	C611516	0.904	-0.170
260	650 - UrbanPopulation	New	C650P001	2	E650001	0.828	-0.234
260	608 - Tree	Trend	C608A513	2	C608513	0.563	-0.471
259	605 - Photo	Trend	C605A506	2	C605506	0.891	-0.214
259	602 - PriceTag	Trend	C602A503	2	C602503	1.134	-0.122
258	623 - Wine	Trend	C623A616	2	C623616	1.018	-0.171
256	646 - RugProduction	New	C646P002	2	E646002	1.042	-0.207
250	613 - Logbook	Trend	C613A520	2	C613520	1.082	-0.301
249	655 - Path	New	C655P001	2	E655001	1.181	-0.294
242	605 - Photo	Trend	C605A507	2	C605507	1.079	-0.447
240	666 - Rope	New	P666P001	2	P666001	0.576	-0.817
239	607 - TV	Trend	C607A510	2	C607510	1.051	-0.513
238	602 - PriceTag	Trend	C602A502	2	C602502	0.648	-0.784
234	665 - Coopertest	New	C665P001	2	E665001	0.932	-0.647
231	615 - Candles	Trend	C615A603	2	C615603	0.929	-0.700
231	645 - AirportTimetable	New	C645P001	2	E645001	0.669	-0.889
228	604 - GasGauge	Trend	C604A505	2	C604505	0.918	-0.771
227	605 - Photo	Trend	C605A508	2	C605508	1.018	-0.739
221	624 - BMI	Trend	C624A619	1	C624619	0.766	-0.987

PIAAC Numeracy Items

Item Difficulty (RP67)	Unit ID and Name	Trend Status	PIAAC Item ID	Proficiency Level (RP67)	Other ID	Slope	Difficulty
221	615 - Candles	Trend	C615A602	1	C615602	0.760	-0.995
217	618 - SixPack1	Trend	C618A607	1	C618607	0.690	-1.115
195	640 - Odometer	New	P640P001	1	P640001	0.909	-1.373
185	614 - Watch	Trend	C614A601	1	C614601	0.808	-1.608
179	635 - ParkingMap	New	C635P001	1	E635001	1.021	-1.615
168	602 - PriceTag	Trend	C602A501	Below 1	C602501	0.678	-2.015
155	600 - Electionresults	Trend	C600AC04	Below 1	C600C04	0.799	-2.160
129	601 - Bottles	Trend	C601AC06	Below 1	C601C06	0.583	-2.827

PIAAC Literacy Items

Item Difficulty (RP67)	Unit ID and Name	Trend Status	PIAAC Item ID	Proficiency Level (RP67)	Other ID	Slope	Difficulty
376	323 - Library Search	New	C323P005	5	E323005	0.967	1.968
374	329 - Work-related Stress	New	C329P003	4	E329003	1.412	2.079
372	306 - CANCO	Trend	C306B111	4	D306111	0.851	1.838
371	308 - Baltic Stock Market	Trend	C308A116	4	C308116	0.735	1.743
359	317 - Apples	New	P317P001	4	P317001	0.782	1.000
350	327 - Summer Streets	New	C327P004	4	E327004	1.132	1.552
349	329 - Work-related Stress	New	C329P002	4	E329002	0.812	1.392
348	323 - Library Search	New	C323P002	4	E323002	1.319	1.568
347	324 - Milk Label	New	P324P002	4	P324002	1.027	1.465
337	308 - Baltic Stock Market	Trend	C308A118	4	C308118	1.009	1.260
329	309 - Generic Medicines	Trend	C309A322	4	C309322	0.776	0.994
329	323 - Library Search	New	C323P004	4	E323004	1.462	1.236
324	313 - International Calls	Trend	C313A410	3	C313410	1.280	1.112
320	327 - Summer Streets	New	C327P003	3	E327003	0.972	0.937
318	315 - Distances-Mexican Cities	Trend	C315B512	3	D315512	0.758	0.766
316	318 - Civil Engineering	New	C318P003	3	E318003	1.250	0.957
315	313 - International Calls	Trend	C313A411	3	C313411	1.516	0.984
312	310 - Memory Training	Trend	C310A407	3	C310407	1.246	0.881
312	324 - Milk Label	New	P324P003	3	P324003	0.892	0.740
309	305 - TMN AntiTheft	Trend	C305A218	3	C305218	1.077	0.764
306	327 - Summer Streets	New	C327P002	3	E327002	0.897	0.632
304	304 - Contact Employer	Trend	C304B711	3	D304711	0.964	0.892
303	318 - Civil Engineering	New	C318P001	3	E318001	1.246	0.703
298	327 - Summer Streets	New	C327P001	3	E327001	0.919	0.492
297	308 - Baltic Stock Market	Trend	C308A119	3	C308119	1.285	0.614
294	322 - Lakeside Fun Run	New	C322P003	3	E322003	1.069	0.478
293	322 - Lakeside Fun Run	New	C322P004	3	E322004	1.442	0.575
289	323 - Library Search	New	C323P003	3	E323003	1.338	0.466
288	307 - MEDCO Aspirin	Trend	C307B402	3	D307402	1.074	0.367
286	320 - Discussion forum	New	C320P003	3	E320003	1.446	0.437
286	313 - International Calls	Trend	C313A413	3	C313413	1.126	0.355
286	304 - Contact Employer	Trend	C304B710	3	D304710	1.722	0.476
285	320 - Discussion forum	New	C320P004	3	E320004	1.338	0.399
283	322 - Lakeside Fun Run	New	C322P001	3	E322001	0.935	0.231
281	320 - Discussion forum	New	C320P001	3	E320001	1.746	0.393
279	308 - Baltic Stock Market	Trend	C308A121	3	C308121	1.296	0.266
272	310 - Memory Training	Trend	C310A406	2	C310406	1.539	0.200
272	309 - Generic Medicines	Trend	C309A319	2	C309319	1.168	0.114
272	313 - International Calls	Trend	C313A414	2	C313414	1.115	0.096

PIAAC Literacy Items

Item Difficulty (RP67)	Unit ID and Name	Trend Status	PIAAC Item ID	Proficiency Level (RP67)	Other ID	Slope	Difficulty
265	317 - Apples	New	P317P003	2	P317003	0.935	-0.121
262	317 - Apples	New	P317P002	2	P317002	1.017	-0.132
260	305 - TMN AntiTheft	Trend	C305A215	2	C305215	1.116	-0.139
257	313 - International Calls	Trend	C313A412	2	C313412	0.926	-0.270
254	308 - Baltic Stock Market	Trend	C308A120	2	C308120	1.270	-0.202
251	321 - Internet Poll	New	C321P001	2	E321001	1.041	-0.329
244	306 - CANCO	Trend	C306B110	2	D306110	1.241	-0.395
244	322 - Lakeside Fun Run	New	C322P005	2	E322005	1.040	-0.465
240	322 - Lakeside Fun Run	New	C322P002	2	E322002	0.858	-0.616
239	308 - Baltic Stock Market	Trend	C308A117	2	C308117	1.088	-0.534
239	309 - Generic Medicines	Trend	C309A320	2	C309320	1.075	-0.549
238	321 - Internet Poll	New	C321P002	2	E321002	0.519	-0.968
219	309 - Generic Medicines	Trend	C309A321	1	C309321	0.984	-0.955
207	330 - Guadeloupe	New	P330P001	1	P330001	0.779	-1.294
201	311 - Dutch Women	Trend	C311B701	1	D311701	0.718	-1.436
169	307 - MEDCO Aspirin	Trend	C30B7401	Below 1	D307401	0.996	-1.883
162	302 - Election Results	Trend	C302BC02	Below 1	D302C02	0.514	-2.411
136	300 - Employment Ad	Trend	C300AC02	Below 1	C300C02	0.785	-2.614
75	301 - SGIH	Trend	C301AC05	Below 1	C301C05	0.502	-4.051

Appendix 2: Contrast Coding Used in Conditioning

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
A_D01a1	14	General - Interview month (Derived by CAPI)	-1	Missing	0000000000001
A_D01a1			1	January	0000000000000
A_D01a1			2	February	1000000000000
A_D01a1			3	March	0100000000000
A_D01a1			4	April	0010000000000
A_D01a1			5	May	0001000000000
A_D01a1			6	June	0000100000000
A_D01a1			7	July	0000010000000
A_D01a1			8	August	0000001000000
A_D01a1			9	September	0000000100000
A_D01a1			10	October	0000000010000
A_D01a1			11	November	0000000001000
A_D01a1			12	Dember	0000000000100
A_D01a1			96	Valid skip	0000000000010
B_D12h	7	Activities - Last year - Number of learning activi	-1	Missing	000001
B_D12h			1	Respondent reported	000000
B_D12h			2	Respondent reported	100000
B_D12h			3	Respondent reported	010000
B_D12h			4	Respondent reported	001000
B_D12h			5	Information on learn	000100
B_D12h			6	Valid skip	000010
B_Q01a	18	Education - Highest qualification - Level	-1	Missing	00000000000000001
B_Q01a			1	No formal qualificat	00000000000000000
B_Q01a			2	ISCED 1	10000000000000000
B_Q01a			3	ISCED 2	01000000000000000
B_Q01a			4	ISCED 3C shorter tha	00100000000000000
B_Q01a			5	ISCED 3C 2 years or	00010000000000000
B_Q01a			6	ISCED 3A-B	00001000000000000
B_Q01a			7	ISCED 3 (without dis	00000100000000000
B_Q01a			8	ISCED 4C	00000010000000000
B_Q01a			9	ISCED 4A-B	00000001000000000
B_Q01a			10	ISCED 4 (without dis	00000000100000000
B_Q01a			11	ISCED 5B	00000000010000000
B_Q01a			12	ISCED 5A, bachelor d	00000000001000000
B_Q01a			13	ISCED 5A, master deg	00000000000100000
B_Q01a			14	ISCED 6	00000000000010000
B_Q01a			15	Foreign qualificatio	00000000000001000
B_Q01a			16	ISCED 5A bachelor de	00000000000000100
B_Q01a			96	Valid skip	00000000000000010
B_Q01a3	17	Education - Highest qualification - Level of forei	-1	Missing	0000000000000001
B_Q01a3			1	No formal qualificat	0000000000000000
B_Q01a3			2	ISCED 1	1000000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a3			3	ISCED 2	0100000000000000
B_Q01a3			4	ISCED 3C shorter tha	0010000000000000
B_Q01a3			5	ISCED 3C 2 years or	0001000000000000
B_Q01a3			6	ISCED 3A-B	0000100000000000
B_Q01a3			7	ISCED 3 (without dis	0000010000000000
B_Q01a3			8	ISCED 4C	0000001000000000
B_Q01a3			9	ISCED 4A-B	0000000100000000
B_Q01a3			10	ISCED 4 (without dis	0000000010000000
B_Q01a3			11	ISCED 5B	0000000001000000
B_Q01a3			12	ISCED 5A, bachelor d	0000000000100000
B_Q01a3			13	ISCED 5A, master deg	0000000000010000
B_Q01a3			14	ISCED 6	0000000000001000
B_Q01a3			15	ISCED 5A bachelor de	0000000000000100
B_Q01a3			96	Valid skip	0000000000000010
B_Q01b	11	Education - Highest qualification - Area of study	-1	Missing	0000000001
B_Q01b			1	General programmes	0000000000
B_Q01b			2	Teacher training and	1000000000
B_Q01b			3	Humanities, language	0100000000
B_Q01b			4	Social sciences, bus	0010000000
B_Q01b			5	Science, mathematics	0001000000
B_Q01b			6	Engineering, manufac	0000100000
B_Q01b			7	Agriculture and vete	0000010000
B_Q01b			8	Health and welfare	0000001000
B_Q01b			9	Services	0000000100
B_Q01b			96	Valid skip	0000000010
B_Q01d	14	Education - Highest qualification - Month of finis	-1	Missing	0000000000001
B_Q01d			1	January	0000000000000
B_Q01d			2	February	1000000000000
B_Q01d			3	March	0100000000000
B_Q01d			4	April	0010000000000
B_Q01d			5	May	0001000000000
B_Q01d			6	June	0000100000000
B_Q01d			7	July	0000010000000
B_Q01d			8	August	0000001000000
B_Q01d			9	September	0000000100000
B_Q01d			10	October	0000000010000
B_Q01d			11	November	0000000001000
B_Q01d			12	Dember	0000000000100
B_Q01d			96	Valid skip	0000000000010
B_Q02a	4	Education - Current qualification	-1	Missing	001
B_Q02a			1	Yes	000
B_Q02a			2	No	100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02a			6	Valid skip	010
B_Q02b	16	Education - Current qualification - Level	-1	Missing	000000000000001
B_Q02b			1	ISCED 1	000000000000000
B_Q02b			2	ISCED 2	100000000000000
B_Q02b			3	ISCED 3C shorter tha	010000000000000
B_Q02b			4	ISCED 3C 2 years or	001000000000000
B_Q02b			5	ISCED 3A-B	000100000000000
B_Q02b			6	ISCED 3 (without dis	000010000000000
B_Q02b			7	ISCED 4C	000001000000000
B_Q02b			8	ISCED 4A-B	000000100000000
B_Q02b			9	ISCED 4 (without dis	000000010000000
B_Q02b			10	ISCED 5B	000000001000000
B_Q02b			11	ISCED 5A, bachelor d	000000000100000
B_Q02b			12	ISCED 5A, master deg	000000000010000
B_Q02b			13	ISCED 6	000000000001000
B_Q02b			14	ISCED 5A bachelor de	000000000000100
B_Q02b			96	Valid skip	000000000000010
B_Q02c	11	Education - Current qualification - Area of study	-1	Missing	0000000001
B_Q02c			1	General programmes	0000000000
B_Q02c			2	Teacher training and	1000000000
B_Q02c			3	Humanities, language	0100000000
B_Q02c			4	Social sciences, bus	0010000000
B_Q02c			5	Science, mathematics	0001000000
B_Q02c			6	Engineering, manufac	0000100000
B_Q02c			7	Agriculture and vete	0000010000
B_Q02c			8	Health and welfare	0000001000
B_Q02c			9	Services	0000000100
B_Q02c			96	Valid skip	0000000010
B_Q03a	4	Education - Uncompleted qualification	-1	Missing	001
B_Q03a			1	Yes	000
B_Q03a			2	No	100
B_Q03a			6	Valid skip	010
B_Q03b	16	Education - Uncompleted qualification - Level	-1	Missing	000000000000001
B_Q03b			1	ISCED 1	000000000000000
B_Q03b			2	ISCED 2	100000000000000
B_Q03b			3	ISCED 3C shorter tha	010000000000000
B_Q03b			4	ISCED 3C 2 years or	001000000000000
B_Q03b			5	ISCED 3A-B	000100000000000
B_Q03b			6	ISCED 3 (without dis	000010000000000
B_Q03b			7	ISCED 4C	000001000000000
B_Q03b			8	ISCED 4A-B	000000100000000
B_Q03b			9	ISCED 4 (without dis	000000010000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q03b			10	ISCED 5B	000000001000000
B_Q03b			11	ISCED 5A, bachelor d	000000000100000
B_Q03b			12	ISCED 5A, master deg	000000000010000
B_Q03b			13	ISCED 6	000000000001000
B_Q03b			14	ISCED 5A bachelor de	000000000000100
B_Q03b			96	Valid skip	000000000000010
B_Q03d	14	Education - Uncompleted qualification - Month of d	-1	Missing	0000000000001
B_Q03d			1	January	0000000000000
B_Q03d			2	February	1000000000000
B_Q03d			3	March	0100000000000
B_Q03d			4	April	0010000000000
B_Q03d			5	May	0001000000000
B_Q03d			6	June	0000100000000
B_Q03d			7	July	0000010000000
B_Q03d			8	August	0000001000000
B_Q03d			9	September	0000000100000
B_Q03d			10	October	0000000010000
B_Q03d			11	November	0000000001000
B_Q03d			12	Dember	0000000000100
B_Q03d			96	Valid skip	0000000000010
B_Q04a	4	Education - Formal qualification	-1	Missing	001
B_Q04a			1	Yes	000
B_Q04a			2	No	100
B_Q04a			6	Valid skip	010
B_Q04b	11	Education - Formal qualification - How many qualif	-1	Missing	0000000001
B_Q04b			1	1 qualification	0000000000
B_Q04b			2	2 qualifications	1000000000
B_Q04b			3	3 qualifications	0100000000
B_Q04b			4	4 qualifications	0010000000
B_Q04b			5	5 qualifications	0001000000
B_Q04b			6	6 qualifications	0000100000
B_Q04b			7	7 qualifications	0000010000
B_Q04b			8	8 qualifications	0000001000
B_Q04b			9	9 qualifications	0000000100
B_Q04b			96	Valid skip	0000000010
B_Q05a	16	Education - Formal qualification - Level	-1	Missing	000000000000001
B_Q05a			1	ISCED 1	000000000000000
B_Q05a			2	ISCED 2	100000000000000
B_Q05a			3	ISCED 3C shorter tha	010000000000000
B_Q05a			4	ISCED 3C 2 years or	001000000000000
B_Q05a			5	ISCED 3A-B	000100000000000
B_Q05a			6	ISCED 3 (without dis	000010000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q05a			7	ISCED 4C	000001000000000
B_Q05a			8	ISCED 4A-B	000000100000000
B_Q05a			9	ISCED 4 (without dis	000000010000000
B_Q05a			10	ISCED 5B	000000001000000
B_Q05a			11	ISCED 5A, bachelor d	000000000100000
B_Q05a			12	ISCED 5A, master deg	000000000010000
B_Q05a			13	ISCED 6	000000000001000
B_Q05a			14	ISCED 5A bachelor de	000000000000100
B_Q05a			96	Valid skip	000000000000010
B_Q05b	11	Education - Formal qualification - Area of study	-1	Missing	0000000001
B_Q05b			1	General programmes	0000000000
B_Q05b			2	Teacher training and	1000000000
B_Q05b			3	Humanities, language	0100000000
B_Q05b			4	Social sciences, bus	0010000000
B_Q05b			5	Science, mathematics	0001000000
B_Q05b			6	Engineering, manufac	0000100000
B_Q05b			7	Agriculture and vete	0000010000
B_Q05b			8	Health and welfare	0000001000
B_Q05b			9	Services	0000000100
B_Q05b			96	Valid skip	0000000010
B_Q05c	4	Education - Formal qualification - Reason job rela	-1	Missing	001
B_Q05c			1	Yes	000
B_Q05c			2	No	100
B_Q05c			6	Valid skip	010
B_Q10a	4	Education - Formal qualification - Employed	-1	Missing	001
B_Q10a			1	Yes	000
B_Q10a			2	No	100
B_Q10a			6	Valid skip	010
B_Q10b	6	Education - Formal qualification - Employed - Work	-1	Missing	00001
B_Q10b			1	Only during working	00000
B_Q10b			2	Mostly during workin	10000
B_Q10b			3	Mostly outside worki	01000
B_Q10b			4	Only outside working	00100
B_Q10b			6	Valid skip	00010
B_Q10c	6	Education - Formal qualification - Employed - Usef	-1	Missing	00001
B_Q10c			1	Not useful at all	00000
B_Q10c			2	Somewhat useful	10000
B_Q10c			3	Moderately useful	01000
B_Q10c			4	Very useful	00100
B_Q10c			6	Valid skip	00010
B_Q11	7	Education - Formal qualification - Grant from empl	-1	Missing	000001
B_Q11			1	Yes, totally	000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q11			2	Yes, partly	100000
B_Q11			3	No, not at all	010000
B_Q11			4	There were no such c	001000
B_Q11			5	No employer or prosp	000100
B_Q11			6	Valid skip	000010
B_Q12a	4	Activities - Last year - Open or distance edu	-1	Missing	001
B_Q12a			1	Yes	000
B_Q12a			2	No	100
B_Q12a			6	Valid skip	010
B_Q12c	4	Activities - Last year - On the job training	-1	Missing	001
B_Q12c			1	Yes	000
B_Q12c			2	No	100
B_Q12c			6	Valid skip	010
B_Q12e	4	Activities - Last year - Seminars or workshops	-1	Missing	001
B_Q12e			1	Yes	000
B_Q12e			2	No	100
B_Q12e			6	Valid skip	010
B_Q12g	4	Activities - Last year - Private lessons	-1	Missing	001
B_Q12g			1	Yes	000
B_Q12g			2	No	100
B_Q12g			6	Valid skip	010
B_Q13	6	Activities - Last year - Activity specified	-1	Missing	00001
B_Q13			1	A course conducted t	00000
B_Q13			2	An organised session	10000
B_Q13			3	A seminar or worksho	01000
B_Q13			4	Other kind of course	00100
B_Q13			6	Valid skip	00010
B_Q14a	4	Activities - Last year - Job related	-1	Missing	001
B_Q14a			1	Yes	000
B_Q14a			2	No	100
B_Q14a			6	Valid skip	010
B_Q14b	10	Activities - Last year - Reason for participating	-1	Missing	000000001
B_Q14b			1	To do my job better	000000000
B_Q14b			2	To be less likely to	100000000
B_Q14b			3	To increase my possi	010000000
B_Q14b			4	To start my own busi	001000000
B_Q14b			5	I was obliged to par	000100000
B_Q14b			6	To increase my knowl	000010000
B_Q14b			7	To obtain a certific	000001000
B_Q14b			8	Other	000000100
B_Q14b			96	Valid skip	000000010
B_Q15a	4	Activities - Last year - Employed	-1	Missing	001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q15a			1	Yes	000
B_Q15a			2	No	100
B_Q15a			6	Valid skip	010
B_Q15b	6	Activities - Last year - During working hours	-1	Missing	00001
B_Q15b			1	Only during working	00000
B_Q15b			2	Mostly during workin	10000
B_Q15b			3	Mostly outside worki	01000
B_Q15b			4	Only outside working	00100
B_Q15b			6	Valid skip	00010
B_Q15c	6	Activities - Last year - Useful for job	-1	Missing	00001
B_Q15c			1	Not useful at all	00000
B_Q15c			2	Somewhat useful	10000
B_Q15c			3	Moderately useful	01000
B_Q15c			4	Very useful	00100
B_Q15c			6	Valid skip	00010
B_Q16	7	Activities - Last year - Grant from employer	-1	Missing	000001
B_Q16			1	Yes, totally	000000
B_Q16			2	Yes, partly	100000
B_Q16			3	No, not at all	010000
B_Q16			4	There were no such c	001000
B_Q16			5	No employer or prosp	000100
B_Q16			6	Valid skip	000010
B_Q17	5	Activities - Last year - Time spend - Unit	-1	Missing	0001
B_Q17			1	Weeks	0000
B_Q17			2	Days	1000
B_Q17			3	Hours	0100
B_Q17			6	Valid skip	0010
B_Q20b	7	Activities - Last year - Time spend for activities	-1	Missing	000001
B_Q20b			1	None of the time	000000
B_Q20b			2	Up to a quarter of t	100000
B_Q20b			3	Up to half of the ti	010000
B_Q20b			4	More than half of th	001000
B_Q20b			5	All of the time	000100
B_Q20b			6	Valid skip	000010
B_Q26a	4	Activities - Last year - Wanted but didn't start	-1	Missing	001
B_Q26a			1	Yes	000
B_Q26a			2	No	100
B_Q26a			6	Valid skip	010
B_Q26b	10	Activities - Last year - Wanted but didn't start -	-1	Missing	000000001
B_Q26b			1	I did not have the p	000000000
B_Q26b			2	Education or trainin	100000000
B_Q26b			3	Education or trais s	010000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q26b			4	I was too busy at wo	001000000
B_Q26b			5	The course or progra	000100000
B_Q26b			6	I did not have time	000010000
B_Q26b			7	Something unexpected	000001000
B_Q26b			8	Other	000000100
B_Q26b			96	Valid skip	000000010
C_D04	5	Current status/work history - Last month - Active	-1	Missing	0001
C_D04			1	Yes	0000
C_D04			2	No	1000
C_D04			3	Not known	0100
C_D04			6	Valid skip	0010
C_D05	6	Current status/work history - Employment status (D	-1	Missing	00001
C_D05			1	Employed	00000
C_D05			2	Unemployed	10000
C_D05			3	Out of the labour fo	01000
C_D05			4	Not known	00100
C_D05			6	Valid skip	00010
C_D06	7	Current status/work history - Current - Paid job o	-1	Missing	000001
C_D06			1	Yes, paid work one j	000000
C_D06			2	Yes, paid work more	100000
C_D06			3	Yes, unpaid work for	010000
C_D06			4	No	001000
C_D06			5	Not known	000100
C_D06			6	Valid skip	000010
C_D08c	4	Current status/work history - Left work in past 5	-1	Missing	001
C_D08c			1	Yes	000
C_D08c			2	No or unknown	100
C_D08c			6	Valid skip	010
C_D09	7	Current status/work history - Work experience (DER	-1	Missing	000001
C_D09			1	Currently working (p	000000
C_D09			2	Recent work experien	100000
C_D09			3	Left paid work longe	010000
C_D09			4	No work experience	001000
C_D09			5	Status unknown	000100
C_D09			6	Valid skip	000010
C_Q01a	4	Current status/work history - Last week - Paid wor	-1	Missing	001
C_Q01a			1	Yes	000
C_Q01a			2	No	100
C_Q01a			6	Valid skip	010
C_Q01b	4	Current status/work history - Last week - Away fro	-1	Missing	001
C_Q01b			1	Yes	000
C_Q01b			2	No	100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
C_Q01b			6	Valid skip	010
C_Q01c	4	Current status/work history - Last week - Unpaid w	-1	Missing	001
C_Q01c			1	Yes	000
C_Q01c			2	No	100
C_Q01c			6	Valid skip	010
C_Q02a	4	Current status/work history - Last month - Looking	-1	Missing	001
C_Q02a			1	Yes	000
C_Q02a			2	No	100
C_Q02a			6	Valid skip	010
C_Q02b	4	Current status/work history - Last month - Waiting	-1	Missing	001
C_Q02b			1	Yes	000
C_Q02b			2	No	100
C_Q02b			6	Valid skip	010
C_Q02c	4	Current status/work history - Last month - Waiting	-1	Missing	001
C_Q02c			1	Within three months	000
C_Q02c			2	In more than three m	100
C_Q02c			6	Valid skip	010
C_Q03_01	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03_01			1	Marked	000
C_Q03_01			2	Not marked	100
C_Q03_01			6	Valid skip	010
C_Q03_02	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03_02			1	Marked	000
C_Q03_02			2	Not marked	100
C_Q03_02			6	Valid skip	010
C_Q03_03	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03_03			1	Marked	000
C_Q03_03			2	Not marked	100
C_Q03_03			6	Valid skip	010
C_Q03_04	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03_04			1	Marked	000
C_Q03_04			2	Not marked	100
C_Q03_04			6	Valid skip	010
C_Q03_05	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03_05			1	Marked	000
C_Q03_05			2	Not marked	100
C_Q03_05			6	Valid skip	010
C_Q03_06	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03_06			1	Marked	000
C_Q03_06			2	Not marked	100
C_Q03_06			6	Valid skip	010
C_Q03_07	4	Current status/work history - Last month - Reason	-1	Missing	001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
C_Q03_07			1	Marked	000
C_Q03_07			2	Not marked	100
C_Q03_07			6	Valid skip	010
C_Q03_08	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03_08			1	Marked	000
C_Q03_08			2	Not marked	100
C_Q03_08			6	Valid skip	010
C_Q03_09	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03_09			1	Marked	000
C_Q03_09			2	Not marked	100
C_Q03_09			6	Valid skip	010
C_Q03_10	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03_10			1	Marked	000
C_Q03_10			2	Not marked	100
C_Q03_10			6	Valid skip	010
C_Q04a	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04a			1	Yes	000
C_Q04a			2	No	100
C_Q04a			6	Valid skip	010
C_Q04b	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04b			1	Yes	000
C_Q04b			2	No	100
C_Q04b			6	Valid skip	010
C_Q04c	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04c			1	Yes	000
C_Q04c			2	No	100
C_Q04c			6	Valid skip	010
C_Q04d	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04d			1	Yes	000
C_Q04d			2	No	100
C_Q04d			6	Valid skip	010
C_Q04e	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04e			1	Yes	000
C_Q04e			2	No	100
C_Q04e			6	Valid skip	010
C_Q04f	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04f			1	Yes	000
C_Q04f			2	No	100
C_Q04f			6	Valid skip	010
C_Q04g	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04g			1	Yes	000
C_Q04g			2	No	100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
C_Q04g			6	Valid skip	010
C_Q04h	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04h			1	Yes	000
C_Q04h			2	No	100
C_Q04h			6	Valid skip	010
C_Q04i	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04i			1	Yes	000
C_Q04i			2	No	100
C_Q04i			6	Valid skip	010
C_Q04j	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04j			1	Yes	000
C_Q04j			2	No	100
C_Q04j			6	Valid skip	010
C_Q05	4	Current status/work history - Ability to start job	-1	Missing	001
C_Q05			1	Yes	000
C_Q05			2	No	100
C_Q05			6	Valid skip	010
C_Q06	4	Current status/work history - Last week - Number o	-1	Missing	001
C_Q06			1	One job or business	000
C_Q06			2	More than one job or	100
C_Q06			6	Valid skip	010
C_Q07	12	Current status/work history - Subjective status	-1	Missing	00000000001
C_Q07			1	Full-time employed (	00000000000
C_Q07			2	Part-time employed (	10000000000
C_Q07			3	Unemployed	01000000000
C_Q07			4	Pupil, student	00100000000
C_Q07			5	Apprentice, internsh	00010000000
C_Q07			6	In retirement or ear	00001000000
C_Q07			7	Permanently disabled	00000100000
C_Q07			8	In compulsory milita	00000010000
C_Q07			9	Fulfilling domestic	00000001000
C_Q07			10	Other	00000000100
C_Q07			96	Valid skip	00000000010
C_Q08a	4	Current status/work history - Ever paid work	-1	Missing	001
C_Q08a			1	Yes	000
C_Q08a			2	No	100
C_Q08a			6	Valid skip	010
C_Q08b	4	Current status/work history - Last year - Paid wor	-1	Missing	001
C_Q08b			1	Yes	000
C_Q08b			2	No	100
C_Q08b			6	Valid skip	010
CBA_START	3	Computer-based exercise agreement	-1	Missing	01

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
CBA_START			1	Continue to computer	00
CBA_START			2	Continue to paper ba	10
CBAMOD1	4	CBA MODULE1 BRANCH	-1	Missing	001
CBAMOD1			1	LIT	000
CBAMOD1			2	NUM	100
CBAMOD1			3	PS1	010
CBAMOD1STG1	4	CBA MODULE1 STAGE1 BRANCH	-1	Missing	001
CBAMOD1STG1			1	EASY	000
CBAMOD1STG1			2	MEDIUM	100
CBAMOD1STG1			3	HARD	010
CBAMOD1STG2	5	CBA MODULE1 STAGE2 BRANCH	-1	Missing	0001
CBAMOD1STG2			1	EASY	0000
CBAMOD1STG2			2	MED1	1000
CBAMOD1STG2			3	MED2	0100
CBAMOD1STG2			4	HARD	0010
CBAMOD2	4	CBA MODULE2 BRANCH	-1	Missing	001
CBAMOD2			1	LIT	000
CBAMOD2			2	NUM	100
CBAMOD2			3	PS2	010
CBAMOD2ALT	8	CBA MODULE1&2 BRANCH	-1	Missing	0000001
CBAMOD2ALT			12	LIT-NUM	0000000
CBAMOD2ALT			13	LIT-PS2	1000000
CBAMOD2ALT			21	NUM-LIT	0100000
CBAMOD2ALT			23	NUM-PS2	0010000
CBAMOD2ALT			31	PS1-LIT	0001000
CBAMOD2ALT			32	PS1-NUM	0000100
CBAMOD2ALT			33	PS1-PS2	0000010
CBAMOD2STG1	4	CBA MODULE2 STAGE1 BRANCH	-1	Missing	001
CBAMOD2STG1			1	EASY	000
CBAMOD2STG1			2	MEDIUM	100
CBAMOD2STG1			3	HARD	010
CBAMOD2STG2	5	CBA MODULE2 STAGE2 BRANCH	-1	Missing	0001
CBAMOD2STG2			1	EASY	0000
CBAMOD2STG2			2	MED1	1000
CBAMOD2STG2			3	MED2	0100
CBAMOD2STG2			4	HARD	0010
E	3	Respondent experience with computer (DERIVED BY CA	-1	Missing	01
E			1	Experienced	00
E			2	Not experienced	10
CORESTAGE1_PASS	3	Core Stage 1 status	-1	Missing	01
CORESTAGE1_PASS			1	Passed	00
CORESTAGE1_PASS			29	Not passed	10

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
CORESTAGE2_PASS	3	Final indication of pass/not pass of Core Stage 2	-1	Missing	01
CORESTAGE2_PASS			1	Passed	00
CORESTAGE2_PASS			29	Not passed	10
D_Q03	5	Current work - Economic sector	-1	Missing	0001
D_Q03			1	The private sector (	0000
D_Q03			2	The public sector (f	1000
D_Q03			3	A non-profit organis	0100
D_Q03			6	Valid skip	0010
D_Q04	4	Current work - Employee or self-employed	-1	Missing	001
D_Q04			1	Employee	000
D_Q04			2	Self-employed	100
D_Q04			6	Valid skip	010
D_Q05a3	14	Current work - Start of work for employer - Month	-1	Missing	0000000000001
D_Q05a3			1	January	0000000000000
D_Q05a3			2	February	1000000000000
D_Q05a3			3	March	0100000000000
D_Q05a3			4	April	0010000000000
D_Q05a3			5	May	0001000000000
D_Q05a3			6	June	0000100000000
D_Q05a3			7	July	0000010000000
D_Q05a3			8	August	0000001000000
D_Q05a3			9	September	0000000100000
D_Q05a3			10	October	0000000010000
D_Q05a3			11	November	0000000001000
D_Q05a3			12	Dember	0000000000100
D_Q05a3			96	Valid skip	0000000000010
D_Q05b3	14	Current work - Start of work for business - Month	-1	Missing	0000000000001
D_Q05b3			1	January	0000000000000
D_Q05b3			2	February	1000000000000
D_Q05b3			3	March	0100000000000
D_Q05b3			4	April	0010000000000
D_Q05b3			5	May	0001000000000
D_Q05b3			6	June	0000100000000
D_Q05b3			7	July	0000010000000
D_Q05b3			8	August	0000001000000
D_Q05b3			9	September	0000000100000
D_Q05b3			10	October	0000000010000
D_Q05b3			11	November	0000000001000
D_Q05b3			12	Dember	0000000000100
D_Q05b3			96	Valid skip	0000000000010
D_Q06a	7	Current work - Amount of people working for employ	-1	Missing	000001
D_Q06a			1	1 to 10 people	000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q06a			2	11 to 50 people	100000
D_Q06a			3	51 to 250 people	010000
D_Q06a			4	251 to 1000 people	001000
D_Q06a			5	More than 1000 peopl	000100
D_Q06a			6	Valid skip	000010
D_Q06b	5	Current work - Amount of people working for employ	-1	Missing	0001
D_Q06b			1	Increased	0000
D_Q06b			2	Decreased	1000
D_Q06b			3	Stayed more or less	0100
D_Q06b			6	Valid skip	0010
D_Q06c	4	Current work - Part of a larger organization	-1	Missing	001
D_Q06c			1	Yes	000
D_Q06c			2	No	100
D_Q06c			6	Valid skip	010
D_Q07a	4	Current work - Employees working for you	-1	Missing	001
D_Q07a			1	Yes	000
D_Q07a			2	No	100
D_Q07a			6	Valid skip	010
D_Q07b	7	Current work - Employees working for you - Amount	-1	Missing	000001
D_Q07b			1	1 to 10 people	000000
D_Q07b			2	11 to 50 people	100000
D_Q07b			3	51 to 250 people	010000
D_Q07b			4	251 to 1000 people	001000
D_Q07b			5	More than 1000 peopl	000100
D_Q07b			6	Valid skip	000010
D_Q08a	4	Current work - Managing other employees	-1	Missing	001
D_Q08a			1	Yes	000
D_Q08a			2	No	100
D_Q08a			6	Valid skip	010
D_Q08b	7	Current work - Managing other employees - Amount	-1	Missing	000001
D_Q08b			1	1 to 5 people	000000
D_Q08b			2	6 to 10 people	100000
D_Q08b			3	11 to 24 people	010000
D_Q08b			4	25 to 99 people	001000
D_Q08b			5	100 or more people	000100
D_Q08b			6	Valid skip	000010
D_Q09	8	Current work - Type of contract	-1	Missing	0000001
D_Q09			1	An indefinite contra	0000000
D_Q09			2	A fixed term contrac	1000000
D_Q09			3	A temporary employme	0100000
D_Q09			4	An apprenticeship or	0010000
D_Q09			5	No contract	0001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q09			6	Other	0000100
D_Q09			96	Valid skip	0000010
D_Q11a	7	Current work - Work flexibility - Sequence of task	-1	Missing	000001
D_Q11a			1	Not at all	000000
D_Q11a			2	Very little	100000
D_Q11a			3	To some extent	010000
D_Q11a			4	To a high extent	001000
D_Q11a			5	To a very high exten	000100
D_Q11a			6	Valid skip	000010
D_Q11b	7	Current work - Work flexibility - How to do the wo	-1	Missing	000001
D_Q11b			1	Not at all	000000
D_Q11b			2	Very little	100000
D_Q11b			3	To some extent	010000
D_Q11b			4	To a high extent	001000
D_Q11b			5	To a very high exten	000100
D_Q11b			6	Valid skip	000010
D_Q11c	7	Current work - Work flexibility - Speed of work	-1	Missing	000001
D_Q11c			1	Not at all	000000
D_Q11c			2	Very little	100000
D_Q11c			3	To some extent	010000
D_Q11c			4	To a high extent	001000
D_Q11c			5	To a very high exten	000100
D_Q11c			6	Valid skip	000010
D_Q11d	7	Current work - Work flexibility - Working hours	-1	Missing	000001
D_Q11d			1	Not at all	000000
D_Q11d			2	Very little	100000
D_Q11d			3	To some extent	010000
D_Q11d			4	To a high extent	001000
D_Q11d			5	To a very high exten	000100
D_Q11d			6	Valid skip	000010
D_Q12a	17	Current work - Requirements - Education level	-1	Missing	0000000000000001
D_Q12a			1	No formal qualificat	0000000000000000
D_Q12a			2	ISCED 1	1000000000000000
D_Q12a			3	ISCED 2	0100000000000000
D_Q12a			4	ISCED 3C shorter tha	0010000000000000
D_Q12a			5	ISCED 3C 2 years or	0001000000000000
D_Q12a			6	ISCED 3A-B	0000100000000000
D_Q12a			7	ISCED 3 (without dis	0000010000000000
D_Q12a			8	ISCED 4C	0000001000000000
D_Q12a			9	ISCED 4A-B	0000000100000000
D_Q12a			10	ISCED 4 (without dis	0000000010000000
D_Q12a			11	ISCED 5B	0000000001000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q12a			12	ISCED 5A, bachelor d	0000000000100000
D_Q12a			13	ISCED 5A, master deg	0000000000010000
D_Q12a			14	ISCED 6	0000000000001000
D_Q12a			15	ISCED 5A bachelor de	0000000000000100
D_Q12a			96	Valid skip	0000000000000010
D_Q12b	5	Current work - Requirements - To do the job satisf	-1	Missing	0001
D_Q12b			1	This level is necess	0000
D_Q12b			2	A lower level would	1000
D_Q12b			3	A higher level would	0100
D_Q12b			6	Valid skip	0010
D_Q12c	8	Current work - Requirements - Related work experie	-1	Missing	0000001
D_Q12c			1	None	0000000
D_Q12c			2	Less than 1 month	1000000
D_Q12c			3	1 to 6 months	0100000
D_Q12c			4	7 to 11 months	0010000
D_Q12c			5	1 or 2 years	0001000
D_Q12c			6	3 years or more	0000100
D_Q12c			96	Valid skip	0000010
D_Q13a	7	Current work - Learning - Learning from co-workers	-1	Missing	000001
D_Q13a			1	Never	000000
D_Q13a			2	Less than once a mon	100000
D_Q13a			3	Less than once a wee	010000
D_Q13a			4	At least once a week	001000
D_Q13a			5	Every day	000100
D_Q13a			6	Valid skip	000010
D_Q13b	7	Current work - Learning - Learning-by-doing	-1	Missing	000001
D_Q13b			1	Never	000000
D_Q13b			2	Less than once a mon	100000
D_Q13b			3	Less than once a wee	010000
D_Q13b			4	At least once a week	001000
D_Q13b			5	Every day	000100
D_Q13b			6	Valid skip	000010
D_Q13c	7	Current work - Learning - Keeping up to date	-1	Missing	000001
D_Q13c			1	Never	000000
D_Q13c			2	Less than once a mon	100000
D_Q13c			3	Less than once a wee	010000
D_Q13c			4	At least once a week	001000
D_Q13c			5	Every day	000100
D_Q13c			6	Valid skip	000010
D_Q14	7	Current work - Job satisfaction	-1	Missing	000001
D_Q14			1	Extremely satisfied	000000
D_Q14			2	Satisfied	100000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q14			3	Neither satisfied no	010000
D_Q14			4	Dissatisfied	001000
D_Q14			5	Extremely dissatisfi	000100
D_Q14			6	Valid skip	000010
D_Q16a	10	Current work - Earnings - Salary interval	-1	Missing	000000001
D_Q16a			1	Per hour	000000000
D_Q16a			2	Per day	100000000
D_Q16a			3	Per week	010000000
D_Q16a			4	Per two weeks	001000000
D_Q16a			5	Per month	000100000
D_Q16a			6	Per year	000010000
D_Q16a			7	Piece rate	000001000
D_Q16a			8	I get no salary or w	000000100
D_Q16a			96	Valid skip	000000010
D_Q16c	4	Current work - Earnings - Gross pay in broad categ	-1	Missing	001
D_Q16c			1	Yes	000
D_Q16c			2	No	100
D_Q16c			6	Valid skip	010
D_Q16d1	8	Current work - Earnings - Broad categories - Gross	-1	Missing	0000001
D_Q16d1			1	Less than 10%	0000000
D_Q16d1			2	10% to less than 2	1000000
D_Q16d1			3	25% to less than 5	0100000
D_Q16d1			4	50% to less than 7	0010000
D_Q16d1			5	75% to less than 9	0001000
D_Q16d1			6	90% or more	0000100
D_Q16d1			96	Valid skip	0000010
D_Q16d2	8	Current work - Earnings - Broad categories - Gross	-1	Missing	0000001
D_Q16d2			1	Less than 10%	0000000
D_Q16d2			2	10% to less than 2	1000000
D_Q16d2			3	25% to less than 5	0100000
D_Q16d2			4	50% to less than 7	0010000
D_Q16d2			5	75% to less than 9	0001000
D_Q16d2			6	90% or more	0000100
D_Q16d2			96	Valid skip	0000010
D_Q16d3	8	Current work - Earnings - Broad categories - Gross	-1	Missing	0000001
D_Q16d3			1	Less than 10%	0000000
D_Q16d3			2	10% to less than 2	1000000
D_Q16d3			3	25% to less than 5	0100000
D_Q16d3			4	50% to less than 7	0010000
D_Q16d3			5	75% to less than 9	0001000
D_Q16d3			6	90% or more	0000100
D_Q16d3			96	Valid skip	0000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q16d4	8	Current work - Earnings - Broad categories - Gross	-1	Missing	0000001
D_Q16d4			1	Less than 10%	0000000
D_Q16d4			2	10% to less than 2	1000000
D_Q16d4			3	25% to less than 5	0100000
D_Q16d4			4	50% to less than 7	0010000
D_Q16d4			5	75% to less than 9	0001000
D_Q16d4			6	90% or more	0000100
D_Q16d4			96	Valid skip	0000010
D_Q16d5	8	Current work - Earnings - Broad categories - Gross	-1	Missing	0000001
D_Q16d5			1	Less than 10%	0000000
D_Q16d5			2	10% to less than 2	1000000
D_Q16d5			3	25% to less than 5	0100000
D_Q16d5			4	50% to less than 7	0010000
D_Q16d5			5	75% to less than 9	0001000
D_Q16d5			6	90% or more	0000100
D_Q16d5			96	Valid skip	0000010
D_Q16d6	8	Current work - Earnings - Broad categories - Gross	-1	Missing	0000001
D_Q16d6			1	Less than 10%	0000000
D_Q16d6			2	10% to less than 2	1000000
D_Q16d6			3	25% to less than 5	0100000
D_Q16d6			4	50% to less than 7	0010000
D_Q16d6			5	75% to less than 9	0001000
D_Q16d6			6	90% or more	0000100
D_Q16d6			96	Valid skip	0000010
D_Q17a	4	Current work - Earnings - Additional payments	-1	Missing	001
D_Q17a			1	Yes	000
D_Q17a			2	No	100
D_Q17a			6	Valid skip	010
D_Q17c	4	Current work - Earnings - Additional payments in b	-1	Missing	001
D_Q17c			1	Yes	000
D_Q17c			2	No	100
D_Q17c			6	Valid skip	010
D_Q17d	5	Current work - Earnings - Additional payments - Br	-1	Missing	0001
D_Q17d			1	Less than 5%	0000
D_Q17d			2	5% to less than 10	1000
D_Q17d			3	10% or more	0100
D_Q17d			6	Valid skip	0010
D_Q18b	4	Current work - Earnings - Total earnings broad cat	-1	Missing	001
D_Q18b			1	Yes	000
D_Q18b			2	No	100
D_Q18b			6	Valid skip	010
D_Q18c1	8	Current work - Earnings - Broad categories - Total	-1	Missing	0000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q18c1			1	Less than 10%	0000000
D_Q18c1			2	10% to less than 2	1000000
D_Q18c1			3	25% to less than 5	0100000
D_Q18c1			4	50% to less than 7	0010000
D_Q18c1			5	75% to less than 9	0001000
D_Q18c1			6	90% or more	0000100
D_Q18c1			96	Valid skip	0000010
D_Q18c2	8	Current work - Earnings - Broad categories - Total	-1	Missing	0000001
D_Q18c2			1	Less than 10%	0000000
D_Q18c2			2	10% to less than 2	1000000
D_Q18c2			3	25% to less than 5	0100000
D_Q18c2			4	50% to less than 7	0010000
D_Q18c2			5	75% to less than 9	0001000
D_Q18c2			6	90% or more	0000100
D_Q18c2			96	Valid skip	0000010
E_Q03	5	Last job - Economic sector	-1	Missing	0001
E_Q03			1	The private sector (	0000
E_Q03			2	The public sector (f	1000
E_Q03			3	A non-profit organis	0100
E_Q03			6	Valid skip	0010
E_Q04	4	Last job - Employee or self-employed	-1	Missing	001
E_Q04			1	Employee	000
E_Q04			2	Self-employed	100
E_Q04			6	Valid skip	010
E_Q06	7	Last job - Amount of people working for employer	-1	Missing	000001
E_Q06			1	1 to 10 people	000000
E_Q06			2	11 to 50 people	100000
E_Q06			3	51 to 250 people	010000
E_Q06			4	251 to 1000 people	001000
E_Q06			5	More than 1000 peopl	000100
E_Q06			6	Valid skip	000010
E_Q07a	4	Last job - Employees working for you	-1	Missing	001
E_Q07a			1	Yes	000
E_Q07a			2	No	100
E_Q07a			6	Valid skip	010
E_Q07b	7	Last job - Employees working for you - Amount	-1	Missing	000001
E_Q07b			1	1 to 10 people	000000
E_Q07b			2	11 to 50 people	100000
E_Q07b			3	51 to 250 people	010000
E_Q07b			4	251 to 1000 people	001000
E_Q07b			5	More than 1000 peopl	000100
E_Q07b			6	Valid skip	000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
E_Q08	8	Last job - Type of contract	-1	Missing	0000001
E_Q08			1	An indefinite contra	0000000
E_Q08			2	A fixed term contrac	1000000
E_Q08			3	A temporary employme	0100000
E_Q08			4	An apprenticeship or	0010000
E_Q08			5	No contract	0001000
E_Q08			6	Other	0000100
E_Q08			96	Valid skip	0000010
E_Q10	12	Last job - Reason for end of job	-1	Missing	00000000001
E_Q10			1	I was dismissed	00000000000
E_Q10			2	I was made redundant	10000000000
E_Q10			3	It was a temporary j	01000000000
E_Q10			4	I resigned	00100000000
E_Q10			5	I gave up work for h	00010000000
E_Q10			6	I took early retirem	00001000000
E_Q10			7	I retired (at or aft	00000100000
E_Q10			8	I gave up work becau	00000010000
E_Q10			9	I gave up work in or	00000001000
E_Q10			10	I left for some othe	00000000100
E_Q10			96	Valid skip	00000000010
EDLEVEL3	4	Educational level of the respondent (DERIVED BY CA	-1	Missing	001
EDLEVEL3			1	Low	000
EDLEVEL3			2	Medium	100
EDLEVEL3			3	High	010
ETSAGEG5	13	Age groups in equal 5-year intervals from 16-65	-1	Missing	000000000001
ETSAGEG5			1	Age 16-20	000000000000
ETSAGEG5			2	Age 21-25	100000000000
ETSAGEG5			3	Age 26-30	010000000000
ETSAGEG5			4	Age 31-35	001000000000
ETSAGEG5			5	Age 36-40	000100000000
ETSAGEG5			6	Age 41-45	000010000000
ETSAGEG5			7	Age 46-50	000001000000
ETSAGEG5			8	Age 51-55	000000100000
ETSAGEG5			9	Age 56-60	000000010000
ETSAGEG5			10	Age 61-65	000000001000
ETSAGEG5			94	Age <16	000000000100
ETSAGEG5			95	Age >65	000000000010
F_Q01b	7	Skill use work - Time cooperating with co-workers	-1	Missing	000001
F_Q01b			1	None of the time	000000
F_Q01b			2	Up to a quarter of t	100000
F_Q01b			3	Up to half of the ti	010000
F_Q01b			4	More than half of th	001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
F_Q01b			5	All of the time	000100
F_Q01b			6	Valid skip	000010
F_Q02a	7	Skill use work - How often - Sharing work-related	-1	Missing	000001
F_Q02a			1	Never	000000
F_Q02a			2	Less than once a mon	100000
F_Q02a			3	Less than once a wee	010000
F_Q02a			4	At least once a week	001000
F_Q02a			5	Every day	000100
F_Q02a			6	Valid skip	000010
F_Q02b	7	Skill use work - How often - Teaching people	-1	Missing	000001
F_Q02b			1	Never	000000
F_Q02b			2	Less than once a mon	100000
F_Q02b			3	Less than once a wee	010000
F_Q02b			4	At least once a week	001000
F_Q02b			5	Every day	000100
F_Q02b			6	Valid skip	000010
F_Q02c	7	Skill use work - How often - Presentations	-1	Missing	000001
F_Q02c			1	Never	000000
F_Q02c			2	Less than once a mon	100000
F_Q02c			3	Less than once a wee	010000
F_Q02c			4	At least once a week	001000
F_Q02c			5	Every day	000100
F_Q02c			6	Valid skip	000010
F_Q02d	7	Skill use work - How often - Selling	-1	Missing	000001
F_Q02d			1	Never	000000
F_Q02d			2	Less than once a mon	100000
F_Q02d			3	Less than once a wee	010000
F_Q02d			4	At least once a week	001000
F_Q02d			5	Every day	000100
F_Q02d			6	Valid skip	000010
F_Q02e	7	Skill use work - How often - Advising people	-1	Missing	000001
F_Q02e			1	Never	000000
F_Q02e			2	Less than once a mon	100000
F_Q02e			3	Less than once a wee	010000
F_Q02e			4	At least once a week	001000
F_Q02e			5	Every day	000100
F_Q02e			6	Valid skip	000010
F_Q03a	7	Skill use work - How often - Planning own activiti	-1	Missing	000001
F_Q03a			1	Never	000000
F_Q03a			2	Less than once a mon	100000
F_Q03a			3	Less than once a wee	010000
F_Q03a			4	At least once a week	001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
F_Q03a			5	Every day	000100
F_Q03a			6	Valid skip	000010
F_Q03b	7	Skill use work - How often - Planning others activ	-1	Missing	000001
F_Q03b			1	Never	000000
F_Q03b			2	Less than once a mon	100000
F_Q03b			3	Less than once a wee	010000
F_Q03b			4	At least once a week	001000
F_Q03b			5	Every day	000100
F_Q03b			6	Valid skip	000010
F_Q03c	7	Skill use work - How often - Organising own time	-1	Missing	000001
F_Q03c			1	Never	000000
F_Q03c			2	Less than once a mon	100000
F_Q03c			3	Less than once a wee	010000
F_Q03c			4	At least once a week	001000
F_Q03c			5	Every day	000100
F_Q03c			6	Valid skip	000010
F_Q04a	7	Skill use work - How often - Influencing people	-1	Missing	000001
F_Q04a			1	Never	000000
F_Q04a			2	Less than once a mon	100000
F_Q04a			3	Less than once a wee	010000
F_Q04a			4	At least once a week	001000
F_Q04a			5	Every day	000100
F_Q04a			6	Valid skip	000010
F_Q04b	7	Skill use work - How often - Negotiating with peop	-1	Missing	000001
F_Q04b			1	Never	000000
F_Q04b			2	Less than once a mon	100000
F_Q04b			3	Less than once a wee	010000
F_Q04b			4	At least once a week	001000
F_Q04b			5	Every day	000100
F_Q04b			6	Valid skip	000010
F_Q05a	7	Skill use work - Problem solving - Simple problems	-1	Missing	000001
F_Q05a			1	Never	000000
F_Q05a			2	Less than once a mon	100000
F_Q05a			3	Less than once a wee	010000
F_Q05a			4	At least once a week	001000
F_Q05a			5	Every day	000100
F_Q05a			6	Valid skip	000010
F_Q05b	7	Skill use work - Problem solving - Complex problem	-1	Missing	000001
F_Q05b			1	Never	000000
F_Q05b			2	Less than once a mon	100000
F_Q05b			3	Less than once a wee	010000
F_Q05b			4	At least once a week	001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
F_Q05b			5	Every day	000100
F_Q05b			6	Valid skip	000010
F_Q06b	7	Skill use work - How often - Working physically fo	-1	Missing	000001
F_Q06b			1	Never	000000
F_Q06b			2	Less than once a mon	100000
F_Q06b			3	Less than once a wee	010000
F_Q06b			4	At least once a week	001000
F_Q06b			5	Every day	000100
F_Q06b			6	Valid skip	000010
F_Q06c	7	Skill use work - How often - Using hands or finger	-1	Missing	000001
F_Q06c			1	Never	000000
F_Q06c			2	Less than once a mon	100000
F_Q06c			3	Less than once a wee	010000
F_Q06c			4	At least once a week	001000
F_Q06c			5	Every day	000100
F_Q06c			6	Valid skip	000010
F_Q07a	4	Skill use work - Not challenged enough	-1	Missing	001
F_Q07a			1	Yes	000
F_Q07a			2	No	100
F_Q07a			6	Valid skip	010
F_Q07b	4	Skill use work - Need more training	-1	Missing	001
F_Q07b			1	Yes	000
F_Q07b			2	No	100
F_Q07b			6	Valid skip	010
G_Q01a	7	Skill use work - Literacy - Read directions or ins	-1	Missing	000001
G_Q01a			1	Never	000000
G_Q01a			2	Less than once a mon	100000
G_Q01a			3	Less than once a wee	010000
G_Q01a			4	At least once a week	001000
G_Q01a			5	Every day	000100
G_Q01a			6	Valid skip	000010
G_Q01b	7	Skill use work - Literacy - Read letters memos or	-1	Missing	000001
G_Q01b			1	Never	000000
G_Q01b			2	Less than once a mon	100000
G_Q01b			3	Less than once a wee	010000
G_Q01b			4	At least once a week	001000
G_Q01b			5	Every day	000100
G_Q01b			6	Valid skip	000010
G_Q01c	7	Skill use work - Literacy - Read newspapers or mag	-1	Missing	000001
G_Q01c			1	Never	000000
G_Q01c			2	Less than once a mon	100000
G_Q01c			3	Less than once a wee	010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
G_Q01c			4	At least once a week	001000
G_Q01c			5	Every day	000100
G_Q01c			6	Valid skip	000010
G_Q01d	7	Skill use work - Literacy - Read professional jour	-1	Missing	000001
G_Q01d			1	Never	000000
G_Q01d			2	Less than once a mon	100000
G_Q01d			3	Less than once a wee	010000
G_Q01d			4	At least once a week	001000
G_Q01d			5	Every day	000100
G_Q01d			6	Valid skip	000010
G_Q01e	7	Skill use work - Literacy - Read books	-1	Missing	000001
G_Q01e			1	Never	000000
G_Q01e			2	Less than once a mon	100000
G_Q01e			3	Less than once a wee	010000
G_Q01e			4	At least once a week	001000
G_Q01e			5	Every day	000100
G_Q01e			6	Valid skip	000010
G_Q01f	7	Skill use work - Literacy - Read manuals or refere	-1	Missing	000001
G_Q01f			1	Never	000000
G_Q01f			2	Less than once a mon	100000
G_Q01f			3	Less than once a wee	010000
G_Q01f			4	At least once a week	001000
G_Q01f			5	Every day	000100
G_Q01f			6	Valid skip	000010
G_Q01g	7	Skill use work - Literacy - Read financial stateme	-1	Missing	000001
G_Q01g			1	Never	000000
G_Q01g			2	Less than once a mon	100000
G_Q01g			3	Less than once a wee	010000
G_Q01g			4	At least once a week	001000
G_Q01g			5	Every day	000100
G_Q01g			6	Valid skip	000010
G_Q01h	7	Skill use work - Literacy - Read diagrams maps or	-1	Missing	000001
G_Q01h			1	Never	000000
G_Q01h			2	Less than once a mon	100000
G_Q01h			3	Less than once a wee	010000
G_Q01h			4	At least once a week	001000
G_Q01h			5	Every day	000100
G_Q01h			6	Valid skip	000010
G_Q02a	7	Skill use work - Literacy - Write letters memos or	-1	Missing	000001
G_Q02a			1	Never	000000
G_Q02a			2	Less than once a mon	100000
G_Q02a			3	Less than once a wee	010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
G_Q02a			4	At least once a week	001000
G_Q02a			5	Every day	000100
G_Q02a			6	Valid skip	000010
G_Q02b	7	Skill use work - Literacy - Write articles	-1	Missing	000001
G_Q02b			1	Never	000000
G_Q02b			2	Less than once a mon	100000
G_Q02b			3	Less than once a wee	010000
G_Q02b			4	At least once a week	001000
G_Q02b			5	Every day	000100
G_Q02b			6	Valid skip	000010
G_Q02c	7	Skill use work - Literacy - Write reports	-1	Missing	000001
G_Q02c			1	Never	000000
G_Q02c			2	Less than once a mon	100000
G_Q02c			3	Less than once a wee	010000
G_Q02c			4	At least once a week	001000
G_Q02c			5	Every day	000100
G_Q02c			6	Valid skip	000010
G_Q02d	7	Skill use work - Literacy - Fill in forms	-1	Missing	000001
G_Q02d			1	Never	000000
G_Q02d			2	Less than once a mon	100000
G_Q02d			3	Less than once a wee	010000
G_Q02d			4	At least once a week	001000
G_Q02d			5	Every day	000100
G_Q02d			6	Valid skip	000010
G_Q03b	7	Skill use work - Numeracy - How often - Calculatin	-1	Missing	000001
G_Q03b			1	Never	000000
G_Q03b			2	Less than once a mon	100000
G_Q03b			3	Less than once a wee	010000
G_Q03b			4	At least once a week	001000
G_Q03b			5	Every day	000100
G_Q03b			6	Valid skip	000010
G_Q03c	7	Skill use work - Numeracy - How often - Use or cal	-1	Missing	000001
G_Q03c			1	Never	000000
G_Q03c			2	Less than once a mon	100000
G_Q03c			3	Less than once a wee	010000
G_Q03c			4	At least once a week	001000
G_Q03c			5	Every day	000100
G_Q03c			6	Valid skip	000010
G_Q03d	7	Skill use work - Numeracy - How often - Use a calc	-1	Missing	000001
G_Q03d			1	Never	000000
G_Q03d			2	Less than once a mon	100000
G_Q03d			3	Less than once a wee	010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
G_Q03d			4	At least once a week	001000
G_Q03d			5	Every day	000100
G_Q03d			6	Valid skip	000010
G_Q03f	7	Skill use work - Numeracy - How often - Prepare ch	-1	Missing	000001
G_Q03f			1	Never	000000
G_Q03f			2	Less than once a mon	100000
G_Q03f			3	Less than once a wee	010000
G_Q03f			4	At least once a week	001000
G_Q03f			5	Every day	000100
G_Q03f			6	Valid skip	000010
G_Q03g	7	Skill use work - Numeracy - How often - Use simple	-1	Missing	000001
G_Q03g			1	Never	000000
G_Q03g			2	Less than once a mon	100000
G_Q03g			3	Less than once a wee	010000
G_Q03g			4	At least once a week	001000
G_Q03g			5	Every day	000100
G_Q03g			6	Valid skip	000010
G_Q03h	7	Skill use work - Numeracy - How often - Use advanc	-1	Missing	000001
G_Q03h			1	Never	000000
G_Q03h			2	Less than once a mon	100000
G_Q03h			3	Less than once a wee	010000
G_Q03h			4	At least once a week	001000
G_Q03h			5	Every day	000100
G_Q03h			6	Valid skip	000010
G_Q04	4	Skill use work - ICT - Experience with computer in	-1	Missing	001
G_Q04			1	Yes	000
G_Q04			2	No	100
G_Q04			6	Valid skip	010
G_Q05a	7	Skill use work - ICT - Internet - How often - For	-1	Missing	000001
G_Q05a			1	Never	000000
G_Q05a			2	Less than once a mon	100000
G_Q05a			3	Less than once a wee	010000
G_Q05a			4	At least once a week	001000
G_Q05a			5	Every day	000100
G_Q05a			6	Valid skip	000010
G_Q05c	7	Skill use work - ICT - Internet - How often - Work	-1	Missing	000001
G_Q05c			1	Never	000000
G_Q05c			2	Less than once a mon	100000
G_Q05c			3	Less than once a wee	010000
G_Q05c			4	At least once a week	001000
G_Q05c			5	Every day	000100
G_Q05c			6	Valid skip	000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
G_Q05d	7	Skill use work - ICT - Internet - How often - Cond	-1	Missing	000001
G_Q05d			1	Never	000000
G_Q05d			2	Less than once a mon	100000
G_Q05d			3	Less than once a wee	010000
G_Q05d			4	At least once a week	001000
G_Q05d			5	Every day	000100
G_Q05d			6	Valid skip	000010
G_Q05e	7	Skill use work - ICT - Computer - How often - Spre	-1	Missing	000001
G_Q05e			1	Never	000000
G_Q05e			2	Less than once a mon	100000
G_Q05e			3	Less than once a wee	010000
G_Q05e			4	At least once a week	001000
G_Q05e			5	Every day	000100
G_Q05e			6	Valid skip	000010
G_Q05f	7	Skill use work - ICT - Computer - How often - Word	-1	Missing	000001
G_Q05f			1	Never	000000
G_Q05f			2	Less than once a mon	100000
G_Q05f			3	Less than once a wee	010000
G_Q05f			4	At least once a week	001000
G_Q05f			5	Every day	000100
G_Q05f			6	Valid skip	000010
G_Q05g	7	Skill use work - ICT - Computer - How often - Prog	-1	Missing	000001
G_Q05g			1	Never	000000
G_Q05g			2	Less than once a mon	100000
G_Q05g			3	Less than once a wee	010000
G_Q05g			4	At least once a week	001000
G_Q05g			5	Every day	000100
G_Q05g			6	Valid skip	000010
G_Q05h	7	Skill use work - ICT - Computer - How often - Real	-1	Missing	000001
G_Q05h			1	Never	000000
G_Q05h			2	Less than once a mon	100000
G_Q05h			3	Less than once a wee	010000
G_Q05h			4	At least once a week	001000
G_Q05h			5	Every day	000100
G_Q05h			6	Valid skip	000010
G_Q06	5	Skill use work - ICT - Computer - Level of compute	-1	Missing	0001
G_Q06			1	Straightforward	0000
G_Q06			2	Moderate	1000
G_Q06			3	Complex	0100
G_Q06			6	Valid skip	0010
G_Q07	4	Skill use work - ICT - Computer - Got the skills n	-1	Missing	001
G_Q07			1	Yes	000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
G_Q07			2	No	100
G_Q07			6	Valid skip	010
G_Q08	4	Skill use work - ICT - Computer - Lack of skills a	-1	Missing	001
G_Q08			1	Yes	000
G_Q08			2	No	100
G_Q08			6	Valid skip	010
GENDER_R	3	Person resolved gender from BQ and QC check (deriv	-1	Missing	01
GENDER_R			1	Male	00
GENDER_R			2	Female	10
GQ_FLAG	3	Group quarters structure flag	-1	Missing	01
GQ_FLAG			0	False	00
GQ_FLAG			1	True	10
H_Q01a	7	Skill use everyday life - Literacy - Read directio	-1	Missing	000001
H_Q01a			1	Never	000000
H_Q01a			2	Less than once a mon	100000
H_Q01a			3	Less than once a wee	010000
H_Q01a			4	At least once a week	001000
H_Q01a			5	Every day	000100
H_Q01a			6	Valid skip	000010
H_Q01b	7	Skill use everyday life - Literacy - Read letters	-1	Missing	000001
H_Q01b			1	Never	000000
H_Q01b			2	Less than once a mon	100000
H_Q01b			3	Less than once a wee	010000
H_Q01b			4	At least once a week	001000
H_Q01b			5	Every day	000100
H_Q01b			6	Valid skip	000010
H_Q01c	7	Skill use everyday life - Literacy - Read newspape	-1	Missing	000001
H_Q01c			1	Never	000000
H_Q01c			2	Less than once a mon	100000
H_Q01c			3	Less than once a wee	010000
H_Q01c			4	At least once a week	001000
H_Q01c			5	Every day	000100
H_Q01c			6	Valid skip	000010
H_Q01d	7	Skill use everyday life - Literacy - Read professi	-1	Missing	000001
H_Q01d			1	Never	000000
H_Q01d			2	Less than once a mon	100000
H_Q01d			3	Less than once a wee	010000
H_Q01d			4	At least once a week	001000
H_Q01d			5	Every day	000100
H_Q01d			6	Valid skip	000010
H_Q01e	7	Skill use everyday life - Literacy - Read books	-1	Missing	000001
H_Q01e			1	Never	000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
H_Q01e			2	Less than once a mon	100000
H_Q01e			3	Less than once a wee	010000
H_Q01e			4	At least once a week	001000
H_Q01e			5	Every day	000100
H_Q01e			6	Valid skip	000010
H_Q01f	7	Skill use everyday life - Literacy - Read manuals	-1	Missing	000001
H_Q01f			1	Never	000000
H_Q01f			2	Less than once a mon	100000
H_Q01f			3	Less than once a wee	010000
H_Q01f			4	At least once a week	001000
H_Q01f			5	Every day	000100
H_Q01f			6	Valid skip	000010
H_Q01g	7	Skill use everyday life - Literacy - Read financia	-1	Missing	000001
H_Q01g			1	Never	000000
H_Q01g			2	Less than once a mon	100000
H_Q01g			3	Less than once a wee	010000
H_Q01g			4	At least once a week	001000
H_Q01g			5	Every day	000100
H_Q01g			6	Valid skip	000010
H_Q01h	7	Skill use everyday life - Literacy - Read diagrams	-1	Missing	000001
H_Q01h			1	Never	000000
H_Q01h			2	Less than once a mon	100000
H_Q01h			3	Less than once a wee	010000
H_Q01h			4	At least once a week	001000
H_Q01h			5	Every day	000100
H_Q01h			6	Valid skip	000010
H_Q02a	7	Skill use everyday life - Literacy - Write letters	-1	Missing	000001
H_Q02a			1	Never	000000
H_Q02a			2	Less than once a mon	100000
H_Q02a			3	Less than once a wee	010000
H_Q02a			4	At least once a week	001000
H_Q02a			5	Every day	000100
H_Q02a			6	Valid skip	000010
H_Q02b	7	Skill use everyday life - Literacy - Write article	-1	Missing	000001
H_Q02b			1	Never	000000
H_Q02b			2	Less than once a mon	100000
H_Q02b			3	Less than once a wee	010000
H_Q02b			4	At least once a week	001000
H_Q02b			5	Every day	000100
H_Q02b			6	Valid skip	000010
H_Q02c	7	Skill use everyday life - Literacy - Write reports	-1	Missing	000001
H_Q02c			1	Never	000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
H_Q02c			2	Less than once a mon	100000
H_Q02c			3	Less than once a wee	010000
H_Q02c			4	At least once a week	001000
H_Q02c			5	Every day	000100
H_Q02c			6	Valid skip	000010
H_Q02d	7	Skill use everyday life - Literacy - Fill in forms	-1	Missing	000001
H_Q02d			1	Never	000000
H_Q02d			2	Less than once a mon	100000
H_Q02d			3	Less than once a wee	010000
H_Q02d			4	At least once a week	001000
H_Q02d			5	Every day	000100
H_Q02d			6	Valid skip	000010
H_Q03b	7	Skill use everyday life - Numeracy - How often - C	-1	Missing	000001
H_Q03b			1	Never	000000
H_Q03b			2	Less than once a mon	100000
H_Q03b			3	Less than once a wee	010000
H_Q03b			4	At least once a week	001000
H_Q03b			5	Every day	000100
H_Q03b			6	Valid skip	000010
H_Q03c	7	Skill use everyday life - Numeracy - How often - U	-1	Missing	000001
H_Q03c			1	Never	000000
H_Q03c			2	Less than once a mon	100000
H_Q03c			3	Less than once a wee	010000
H_Q03c			4	At least once a week	001000
H_Q03c			5	Every day	000100
H_Q03c			6	Valid skip	000010
H_Q03d	7	Skill use everyday life - Numeracy - How often - U	-1	Missing	000001
H_Q03d			1	Never	000000
H_Q03d			2	Less than once a mon	100000
H_Q03d			3	Less than once a wee	010000
H_Q03d			4	At least once a week	001000
H_Q03d			5	Every day	000100
H_Q03d			6	Valid skip	000010
H_Q03f	7	Skill use everyday life - Numeracy - How often - P	-1	Missing	000001
H_Q03f			1	Never	000000
H_Q03f			2	Less than once a mon	100000
H_Q03f			3	Less than once a wee	010000
H_Q03f			4	At least once a week	001000
H_Q03f			5	Every day	000100
H_Q03f			6	Valid skip	000010
H_Q03g	7	Skill use everyday life - Numeracy - How often - U	-1	Missing	000001
H_Q03g			1	Never	000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
H_Q03g			2	Less than once a mon	100000
H_Q03g			3	Less than once a wee	010000
H_Q03g			4	At least once a week	001000
H_Q03g			5	Every day	000100
H_Q03g			6	Valid skip	000010
H_Q03h	7	Skill use everyday life - Numeracy - How often - U	-1	Missing	000001
H_Q03h			1	Never	000000
H_Q03h			2	Less than once a mon	100000
H_Q03h			3	Less than once a wee	010000
H_Q03h			4	At least once a week	001000
H_Q03h			5	Every day	000100
H_Q03h			6	Valid skip	000010
H_Q04a	4	Skill use everyday life - ICT - Ever used computer	-1	Missing	001
H_Q04a			1	Yes	000
H_Q04a			2	No	100
H_Q04a			6	Valid skip	010
H_Q04b	4	Skill use everyday life - ICT - Experience with co	-1	Missing	001
H_Q04b			1	Yes	000
H_Q04b			2	No	100
H_Q04b			6	Valid skip	010
H_Q05a	7	Skill use everyday life - ICT - Internet - How oft	-1	Missing	000001
H_Q05a			1	Never	000000
H_Q05a			2	Less than once a mon	100000
H_Q05a			3	Less than once a wee	010000
H_Q05a			4	At least once a week	001000
H_Q05a			5	Every day	000100
H_Q05a			6	Valid skip	000010
H_Q05c	7	Skill use everyday life - ICT - Internet - How oft	-1	Missing	000001
H_Q05c			1	Never	000000
H_Q05c			2	Less than once a mon	100000
H_Q05c			3	Less than once a wee	010000
H_Q05c			4	At least once a week	001000
H_Q05c			5	Every day	000100
H_Q05c			6	Valid skip	000010
H_Q05d	7	Skill use everyday life - ICT - Internet - How oft	-1	Missing	000001
H_Q05d			1	Never	000000
H_Q05d			2	Less than once a mon	100000
H_Q05d			3	Less than once a wee	010000
H_Q05d			4	At least once a week	001000
H_Q05d			5	Every day	000100
H_Q05d			6	Valid skip	000010
H_Q05e	7	Skill use everyday life - ICT - Computer - How oft	-1	Missing	000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
H_Q05e			1	Never	000000
H_Q05e			2	Less than once a mon	100000
H_Q05e			3	Less than once a wee	010000
H_Q05e			4	At least once a week	001000
H_Q05e			5	Every day	000100
H_Q05e			6	Valid skip	000010
H_Q05f	7	Skill use everyday life - ICT - Computer - How oft	-1	Missing	000001
H_Q05f			1	Never	000000
H_Q05f			2	Less than once a mon	100000
H_Q05f			3	Less than once a wee	010000
H_Q05f			4	At least once a week	001000
H_Q05f			5	Every day	000100
H_Q05f			6	Valid skip	000010
H_Q05g	7	Skill use everyday life - ICT - Computer - How oft	-1	Missing	000001
H_Q05g			1	Never	000000
H_Q05g			2	Less than once a mon	100000
H_Q05g			3	Less than once a wee	010000
H_Q05g			4	At least once a week	001000
H_Q05g			5	Every day	000100
H_Q05g			6	Valid skip	000010
H_Q05h	7	Skill use everyday life - ICT - Computer - How oft	-1	Missing	000001
H_Q05h			1	Never	000000
H_Q05h			2	Less than once a mon	100000
H_Q05h			3	Less than once a wee	010000
H_Q05h			4	At least once a week	001000
H_Q05h			5	Every day	000100
H_Q05h			6	Valid skip	000010
HIDD_DU	3	Hidden dwelling unit (DU)	-1	Missing	01
HIDD_DU			0	False	00
HIDD_DU			1	True	10
I_Q04b	7	About yourself - Learning strategies - Relate new	-1	Missing	000001
I_Q04b			1	Not at all	000000
I_Q04b			2	Very little	100000
I_Q04b			3	To some extent	010000
I_Q04b			4	To a high extent	001000
I_Q04b			5	To a very high exten	000100
I_Q04b			6	Valid skip	000010
I_Q04d	7	About yourself - Learning strategies - Like learni	-1	Missing	000001
I_Q04d			1	Not at all	000000
I_Q04d			2	Very little	100000
I_Q04d			3	To some extent	010000
I_Q04d			4	To a high extent	001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
I_Q04d			5	To a very high exten	000100
I_Q04d			6	Valid skip	000010
I_Q04h	7	About yourself - Learning strategies - Attribute s	-1	Missing	000001
I_Q04h			1	Not at all	000000
I_Q04h			2	Very little	100000
I_Q04h			3	To some extent	010000
I_Q04h			4	To a high extent	001000
I_Q04h			5	To a very high exten	000100
I_Q04h			6	Valid skip	000010
I_Q04j	7	About yourself - Learning strategies - Get to the	-1	Missing	000001
I_Q04j			1	Not at all	000000
I_Q04j			2	Very little	100000
I_Q04j			3	To some extent	010000
I_Q04j			4	To a high extent	001000
I_Q04j			5	To a very high exten	000100
I_Q04j			6	Valid skip	000010
I_Q04l	7	About yourself - Learning strategies - Figure out	-1	Missing	000001
I_Q04l			1	Not at all	000000
I_Q04l			2	Very little	100000
I_Q04l			3	To some extent	010000
I_Q04l			4	To a high extent	001000
I_Q04l			5	To a very high exten	000100
I_Q04l			6	Valid skip	000010
I_Q04m	7	About yourself - Learning strategies - Looking for	-1	Missing	000001
I_Q04m			1	Not at all	000000
I_Q04m			2	Very little	100000
I_Q04m			3	To some extent	010000
I_Q04m			4	To a high extent	001000
I_Q04m			5	To a very high exten	000100
I_Q04m			6	Valid skip	000010
I_Q05f	7	About yourself - Cultural engagement - Voluntary w	-1	Missing	000001
I_Q05f			1	Never	000000
I_Q05f			2	Less than once a mon	100000
I_Q05f			3	Less than once a wee	010000
I_Q05f			4	At least once a week	001000
I_Q05f			5	Every day	000100
I_Q05f			6	Valid skip	000010
I_Q06a	7	About yourself - Political efficacy - No influence	-1	Missing	000001
I_Q06a			1	Strongly agree	000000
I_Q06a			2	Agree	100000
I_Q06a			3	Neither agree nor di	010000
I_Q06a			4	Disagree	001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
I_Q06a			5	Strongly disagree	000100
I_Q06a			6	Valid skip	000010
I_Q07a	7	About yourself - Social trust - Trust only few peo	-1	Missing	000001
I_Q07a			1	Strongly agree	000000
I_Q07a			2	Agree	100000
I_Q07a			3	Neither agree nor di	010000
I_Q07a			4	Disagree	001000
I_Q07a			5	Strongly disagree	000100
I_Q07a			6	Valid skip	000010
I_Q07b	7	About yourself - Social trust - Other people take	-1	Missing	000001
I_Q07b			1	Strongly agree	000000
I_Q07b			2	Agree	100000
I_Q07b			3	Neither agree nor di	010000
I_Q07b			4	Disagree	001000
I_Q07b			5	Strongly disagree	000100
I_Q07b			6	Valid skip	000010
I_Q08	7	About yourself - Health - State	-1	Missing	000001
I_Q08			1	Excellent	000000
I_Q08			2	Very good	100000
I_Q08			3	Good	010000
I_Q08			4	Fair	001000
I_Q08			5	Poor	000100
I_Q08			6	Valid skip	000010
ISCED_HF	17	Level of Highest Qualification (Foreign) - Respond	-1	Missing	0000000000000001
ISCED_HF			1	No formal qualificat	0000000000000000
ISCED_HF			2	ISCED 1	1000000000000000
ISCED_HF			3	ISCED 2	0100000000000000
ISCED_HF			4	ISCED 3C shorter tha	0010000000000000
ISCED_HF			5	ISCED 3C 2 years or	0001000000000000
ISCED_HF			6	ISCED 3A-B	0000100000000000
ISCED_HF			7	ISCED 3 (without dis	0000010000000000
ISCED_HF			8	ISCED 4C	0000001000000000
ISCED_HF			9	ISCED 4A-B	0000000100000000
ISCED_HF			10	ISCED 4 (without dis	0000000010000000
ISCED_HF			11	ISCED 5B	0000000001000000
ISCED_HF			12	ISCED 5A, bachelor d	0000000000100000
ISCED_HF			13	ISCED 5A, master deg	0000000000010000
ISCED_HF			14	ISCED 6	0000000000001000
ISCED_HF			15	ISCED 5A bachelor de	0000000000000100
ISCED_HF			96	Valid skip	0000000000000010
J_N05a2	4	Background - More than one language mentioned	-1	Missing	001
J_N05a2			1	Yes	000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_N05a2			2	No	100
J_N05a2			6	Valid skip	010
J_Q02a	4	Background - Living with spouse or partner	-1	Missing	001
J_Q02a			1	Yes	000
J_Q02a			2	No	100
J_Q02a			6	Valid skip	010
J_Q02c	12	Background - Work situation of spouse or partner	-1	Missing	00000000001
J_Q02c			1	Full-time employed (	00000000000
J_Q02c			2	Part-time employed (	10000000000
J_Q02c			3	Unemployed	01000000000
J_Q02c			4	Pupil, student	00100000000
J_Q02c			5	Apprentice, internsh	00010000000
J_Q02c			6	In retirement or ear	00001000000
J_Q02c			7	Permanently disabled	00000100000
J_Q02c			8	In compulsory milita	00000010000
J_Q02c			9	Fulfilling domestic	00000001000
J_Q02c			10	Other	00000000100
J_Q02c			96	Valid skip	00000000010
J_Q03a	4	Background - Children	-1	Missing	001
J_Q03a			1	Yes	000
J_Q03a			2	No	100
J_Q03a			6	Valid skip	010
J_Q03b	27	Background - Number of children	-1	Missing	000000000000000000000000001
J_Q03b			1	1 kid	000000000000000000000000000
J_Q03b			2	2 kids	100000000000000000000000000
J_Q03b			3	3 kids	010000000000000000000000000
J_Q03b			4	4 kids	001000000000000000000000000
J_Q03b			5	5 kids	000100000000000000000000000
J_Q03b			6	6 kids	000010000000000000000000000
J_Q03b			7	7 kids	000001000000000000000000000
J_Q03b			8	8 kids	000000100000000000000000000
J_Q03b			9	9 kids	000000010000000000000000000
J_Q03b			10	10 kids	000000001000000000000000000
J_Q03b			11	11 kids	000000000100000000000000000
J_Q03b			12	12 kids	000000000010000000000000000
J_Q03b			13	13 kids	000000000001000000000000000
J_Q03b			14	14 kids	000000000000100000000000000
J_Q03b			15	15 kids	000000000000010000000000000
J_Q03b			16	16 kids	000000000000001000000000000
J_Q03b			17	17 kids	000000000000000100000000000
J_Q03b			18	18 kids	000000000000000010000000000
J_Q03b			19	19 kids	000000000000000001000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q03b			20	20 kids	0000000000000000010000000
J_Q03b			21	21 kids	0000000000000000001000000
J_Q03b			22	22 kids	0000000000000000000100000
J_Q03b			23	23 kids	0000000000000000000010000
J_Q03b			24	24 kids	0000000000000000000001000
J_Q03b			25	25 kids	0000000000000000000000100
J_Q03b			96	Valid skip	0000000000000000000000010
J_Q04a	4	Background - Born in country	-1	Missing	001
J_Q04a			1	Yes	000
J_Q04a			2	No	100
J_Q04a			6	Valid skip	010
J_Q06a	4	Background - Mother/female guardian - Whether born	-1	Missing	001
J_Q06a			1	Yes	000
J_Q06a			2	No	100
J_Q06a			6	Valid skip	010
J_Q06b	5	Background - Mother/female guardian - Highest leve	-1	Missing	0001
J_Q06b			1	ISCED 1, 2, and 3C s	0000
J_Q06b			2	ISCED 3 (excl 3C sho	1000
J_Q06b			3	ISCED 5 and 6	0100
J_Q06b			6	Valid skip	0010
J_Q06c	5	Background - Mother/female guardian - Occupational	-1	Missing	0001
J_Q06c			1	Yes	0000
J_Q06c			2	No	1000
J_Q06c			3	Not applicable (Moth	0100
J_Q06c			6	Valid skip	0010
J_Q07a	4	Background - Father/male guardian - Whether born i	-1	Missing	001
J_Q07a			1	Yes	000
J_Q07a			2	No	100
J_Q07a			6	Valid skip	010
J_Q07b	5	Background - Father/male guardian - Highest level	-1	Missing	0001
J_Q07b			1	ISCED 1, 2, and 3C s	0000
J_Q07b			2	ISCED 3 (excl 3C sho	1000
J_Q07b			3	ISCED 5 and 6	0100
J_Q07b			6	Valid skip	0010
J_Q07c	5	Background - Father/male guardian - Occupational s	-1	Missing	0001
J_Q07c			1	Yes	0000
J_Q07c			2	No	1000
J_Q07c			3	Not applicable (Moth	0100
J_Q07c			6	Valid skip	0010
J_Q08	8	Background - Number of books at home	-1	Missing	0000001
J_Q08			1	10 books or less	0000000
J_Q08			2	11 to 25 books	1000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q08			3	26 to 100 books	0100000
J_Q08			4	101 to 200 books	0010000
J_Q08			5	201 to 500 books	0001000
J_Q08			6	More than 500 books	0000100
J_Q08			96	Valid skip	0000010
NATIVESPEAKER	4	Respondent is a native speaker (DERIVED BY CAPI)	-1	Missing	001
NATIVESPEAKER			1	Yes	000
NATIVESPEAKER			2	No	100
NATIVESPEAKER			6	Valid skip	010
PAPER	4	PAPER BRANCH	-1	Missing	001
PAPER			1	PP1-LIT	000
PAPER			2	PP2-NUM	100
PAPER			3	Failed PaperCore	010
PBROUTE	6	Paper-Based Routing	-1	Missing	00001
PBROUTE			1	No comp experience	00000
PBROUTE			2	Failed ICTcorestage1	10000
PBROUTE			3	Refused CBA	01000
PBROUTE			4	CBA	00100
PBROUTE			5	Uncategorized	00010
TECHPROB	7	Technical problem flag	-1	Missing	000001
TECHPROB			1	Zip file exists but	000000
TECHPROB			2	Zip file exists but	100000
TECHPROB			3	VM froze/crashed and	010000
TECHPROB			4	VM froze/crashed and	001000
TECHPROB			5	Scripts did not func	000100
TECHPROB			6	Other	000010
VET	4	Actual (sels highest level of education is vocatio	-1	Missing	001
VET			0	False	000
VET			1	True	100
VET			6	Valid skip	010
YEARLYINCPR	7	Categorical yearly income	-1	Missing	000001
YEARLYINCPR			1	Less than 10%	000000
YEARLYINCPR			2	10% to less than 25%	100000
YEARLYINCPR			3	25% to less than 50%	010000
YEARLYINCPR			4	50% to less than 75%	001000
YEARLYINCPR			5	75% to less than 90%	000100
YEARLYINCPR			6	90% or more	000010
ZZ1a	4	Observation module: Presence of additional person	-1	Missing	001
ZZ1a			1	Yes	000
ZZ1a			2	No	100
ZZ1a			6	Valid skip	010
ZZ1b_01	4	Observation module: Assistance in background quest	-1	Missing	001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
ZZ1b_01			1	Yes	000
ZZ1b_01			2	No	100
ZZ1b_01			6	Valid skip	010
ZZ1b_02	4	Observation module: Assistance in skills assessmen	-1	Missing	001
ZZ1b_02			1	Yes	000
ZZ1b_02			2	No	100
ZZ1b_02			6	Valid skip	010
ZZ2	6	Observation module: Respondent understood the ques	-1	Missing	00001
ZZ2			1	Never	00000
ZZ2			2	Almost never	10000
ZZ2			3	Now and then	01000
ZZ2			4	Often	00100
ZZ2			5	Very Often	00010
ZZ3	4	Observation module: Clarification necessary	-1	Missing	001
ZZ3			1	Yes	000
ZZ3			2	No	100
ZZ3			6	Valid skip	010
ZZ4_01	4	Observation module: Respondent held a conversation	-1	Missing	001
ZZ4_01			1	Yes	000
ZZ4_01			2	No	100
ZZ4_01			6	Valid skip	010
ZZ4_02	4	Observation module: Respondent answered a phone ca	-1	Missing	001
ZZ4_02			1	Yes	000
ZZ4_02			2	No	100
ZZ4_02			6	Valid skip	010
ZZ4_03	4	Observation module: Respondent was looking after c	-1	Missing	001
ZZ4_03			1	Yes	000
ZZ4_03			2	No	100
ZZ4_03			6	Valid skip	010
ZZ4_04	4	Observation module: Respondent was undertaking dom	-1	Missing	001
ZZ4_04			1	Yes	000
ZZ4_04			2	No	100
ZZ4_04			6	Valid skip	010
ZZ4_05	4	Observation module: Television, radio, game consol	-1	Missing	001
ZZ4_05			1	Yes	000
ZZ4_05			2	No	100
ZZ4_05			6	Valid skip	010
ZZ4_06	4	Observation module: Respondent was interrupted by	-1	Missing	001
ZZ4_06			1	Yes	000
ZZ4_06			2	No	100
ZZ4_06			6	Valid skip	010
ZZ5	4	Observation module: Assessment taking too long	-1	Missing	001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
ZZ5			1	Yes	000
ZZ5			2	No	100
ZZ5			6	Valid skip	010
ZZ6	9	Observation module: Room of assessment	-1	Missing	00000001
ZZ6			1	Living/dining room	00000000
ZZ6			2	Kitchen	10000000
ZZ6			3	Bedroom	01000000
ZZ6			4	Entrance	00100000
ZZ6			5	Hallway or corridor	00010000
ZZ6			6	Office	00001000
ZZ6			7	Other space in the h	00000100
ZZ6			8	Other space outside	00000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
A_N01UKX	6	Country in which interview conducted	-1	Missing	00001
A_N01UKX			1	England	00000
A_N01UKX			2	Wales	10000
A_N01UKX			3	Scotland	01000
A_N01UKX			4	Northern Ireland	00100
A_N01UKX			6	Valid skip	00010
A_N03a1ca	4	Language - More than one language mentioned	-1	Missing	001
A_N03a1ca			1	Yes	000
A_N03a1ca			2	No	100
A_N03a1ca			6	Valid skip	010
A_Q01BCA1	9	Respondent age range	-1	Missing	00000001
A_Q01BCA1			1	Less than 16 years	00000000
A_Q01BCA1			2	16-24	10000000
A_Q01BCA1			3	25-34	01000000
A_Q01BCA1			4	35-44	00100000
A_Q01BCA1			5	45-54	00010000
A_Q01BCA1			6	55-65	00001000
A_Q01BCA1			7	66 and over	00000100
A_Q01BCA1			96	Valid skip	00000010
A_Q02CA	4	Background - Born in Canada	-1	Missing	001
A_Q02CA			1	Yes	000
A_Q02CA			2	No	100
A_Q02CA			6	Valid skip	010
A_Q03A1CA	15	Language - First learned language	-1	Missing	00000000000001
A_Q03A1CA			1	English	00000000000000
A_Q03A1CA			2	French	10000000000000
A_Q03A1CA			3	Italian	01000000000000
A_Q03A1CA			4	Chinese	00100000000000
A_Q03A1CA			5	German	00010000000000
A_Q03A1CA			6	Portuguese	00001000000000
A_Q03A1CA			7	Polish	00000100000000
A_Q03A1CA			8	Ukrainian	00000010000000
A_Q03A1CA			9	Spanish	00000001000000
A_Q03A1CA			10	Dutch	00000000100000
A_Q03A1CA			11	Punjabi	00000000010000
A_Q03A1CA			12	Greek	00000000001000
A_Q03A1CA			13	Other - specify	00000000000100
A_Q03A1CA			96	Valid skip	00000000000010
A_Q03a2ca	15	Language - Second learned language	-1	Missing	00000000000001
A_Q03a2ca			1	English	00000000000000
A_Q03a2ca			2	French	10000000000000
A_Q03a2ca			3	Italian	01000000000000
A_Q03a2ca			4	Chinese	00100000000000
A_Q03a2ca			5	German	00010000000000
A_Q03a2ca			6	Portuguese	00001000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
A_Q03a2ca			7	Polish	00000100000000
A_Q03a2ca			8	Ukrainian	00000010000000
A_Q03a2ca			9	Spanish	00000001000000
A_Q03a2ca			10	Dutch	00000000100000
A_Q03a2ca			11	Punjabi	00000000010000
A_Q03a2ca			12	Greek	00000000001000
A_Q03a2ca			13	Other - specify	00000000000100
A_Q03a2ca			96	Valid skip	00000000000010
A_Q04bca	15	Language - Language spoken at home	-1	Missing	00000000000001
A_Q04bca			1	English	00000000000000
A_Q04bca			2	French	10000000000000
A_Q04bca			3	Italian	01000000000000
A_Q04bca			4	Chinese	00100000000000
A_Q04bca			5	German	00010000000000
A_Q04bca			6	Portuguese	00001000000000
A_Q04bca			7	Polish	00000100000000
A_Q04bca			8	Ukrainian	00000010000000
A_Q04bca			9	Spanish	00000001000000
A_Q04bca			10	Dutch	00000000100000
A_Q04bca			11	Punjabi	00000000010000
A_Q04bca			12	Greek	00000000001000
A_Q04bca			13	Other - specify	00000000000100
A_Q04bca			96	Valid skip	00000000000010
A_Q04cca	4	Language - Other language spoken at home - Yes/No	-1	Missing	001
A_Q04cca			1	Yes	000
A_Q04cca			2	No	100
A_Q04cca			6	Valid skip	010
A_Q04cca1_01	4	Language - Other language spoken at home - English	-1	Missing	001
A_Q04cca1_01			1	Marked	000
A_Q04cca1_01			2	Not marked	100
A_Q04cca1_01			6	Valid skip	010
A_Q04cca1_02	4	Language - Other language spoken at home - French	-1	Missing	001
A_Q04cca1_02			1	Marked	000
A_Q04cca1_02			2	Not marked	100
A_Q04cca1_02			6	Valid skip	010
A_Q04cca1_03	4	Language - Other language spoken at home - Italian	-1	Missing	001
A_Q04cca1_03			1	Marked	000
A_Q04cca1_03			2	Not marked	100
A_Q04cca1_03			6	Valid skip	010
A_Q04cca1_04	4	Language - Other language spoken at home - Chinese	-1	Missing	001
A_Q04cca1_04			1	Marked	000
A_Q04cca1_04			2	Not marked	100
A_Q04cca1_04			6	Valid skip	010
A_Q04cca1_05	4	Language - Other language spoken at home - German	-1	Missing	001
A_Q04cca1_05			1	Marked	000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
A_Q04cca1_05			2	Not marked	100
A_Q04cca1_05			6	Valid skip	010
A_Q04cca1_06	4	Language - Other language spoken at home - Portugu	-1	Missing	001
A_Q04cca1_06			1	Marked	000
A_Q04cca1_06			2	Not marked	100
A_Q04cca1_06			6	Valid skip	010
A_Q04cca1_07	4	Language - Other language spoken at home - Polish	-1	Missing	001
A_Q04cca1_07			1	Marked	000
A_Q04cca1_07			2	Not marked	100
A_Q04cca1_07			6	Valid skip	010
A_Q04cca1_08	4	Language - Other language spoken at home - Ukraini	-1	Missing	001
A_Q04cca1_08			1	Marked	000
A_Q04cca1_08			2	Not marked	100
A_Q04cca1_08			6	Valid skip	010
A_Q04cca1_09	4	Language - Other language spoken at home - Spanish	-1	Missing	001
A_Q04cca1_09			1	Marked	000
A_Q04cca1_09			2	Not marked	100
A_Q04cca1_09			6	Valid skip	010
A_Q04cca1_10	4	Language - Other language spoken at home - Dutch	-1	Missing	001
A_Q04cca1_10			1	Marked	000
A_Q04cca1_10			2	Not marked	100
A_Q04cca1_10			6	Valid skip	010
A_Q04cca1_11	4	Language - Other language spoken at home - Punjabi	-1	Missing	001
A_Q04cca1_11			1	Marked	000
A_Q04cca1_11			2	Not marked	100
A_Q04cca1_11			6	Valid skip	010
A_Q04cca1_12	4	Language - Other language spoken at home - Greek	-1	Missing	001
A_Q04cca1_12			1	Marked	000
A_Q04cca1_12			2	Not marked	100
A_Q04cca1_12			6	Valid skip	010
A_Q04cca1_13	4	Language - Other language spoken at home - Other-	-1	Missing	001
A_Q04cca1_13			1	Marked	000
A_Q04cca1_13			2	Not marked	100
A_Q04cca1_13			6	Valid skip	010
A_Q04fca	7	Language - Current reading skills in English/Frenc	-1	Missing	000001
A_Q04fca			1	Cannot read this lan	000000
A_Q04fca			2	Poor	100000
A_Q04fca			3	Fair	010000
A_Q04fca			4	Good	001000
A_Q04fca			5	Very good	000100
A_Q04fca			6	Valid skip	000010
A_Q04gca	7	Language - Current writing skills in English/Frenc	-1	Missing	000001
A_Q04gca			1	Cannot write in this	000000
A_Q04gca			2	Poor	100000
A_Q04gca			3	Fair	010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
A_Q04gca			4	Good	001000
A_Q04gca			5	Very good	000100
A_Q04gca			6	Valid skip	000010
A_Q04ica	7	Language - Current reading skills in English/Frenc	-1	Missing	000001
A_Q04ica			1	Cannot read this lan	000000
A_Q04ica			2	Poor	100000
A_Q04ica			3	Fair	010000
A_Q04ica			4	Good	001000
A_Q04ica			5	Very good	000100
A_Q04ica			6	Valid skip	000010
A_Q04jca	7	Language - Current writing skills in English/Frenc	-1	Missing	000001
A_Q04jca			1	Cannot write in this	000000
A_Q04jca			2	Poor	100000
A_Q04jca			3	Fair	010000
A_Q04jca			4	Good	001000
A_Q04jca			5	Very good	000100
A_Q04jca			6	Valid skip	000010
A_Q04lca1	7	Language - Current ability to speak English/French	-1	Missing	000001
A_Q04lca1			1	Cannot speak in this	000000
A_Q04lca1			2	Poor	100000
A_Q04lca1			3	Fair	010000
A_Q04lca1			4	Good	001000
A_Q04lca1			5	Very good	000100
A_Q04lca1			6	Valid skip	000010
A_Q04lca2	7	Language - Current ability to speak English/French	-1	Missing	000001
A_Q04lca2			1	Cannot speak in this	000000
A_Q04lca2			2	Poor	100000
A_Q04lca2			3	Fair	010000
A_Q04lca2			4	Good	001000
A_Q04lca2			5	Very good	000100
A_Q04lca2			6	Valid skip	000010
AA2	5	Respondent Language of Preference - From CMS	-1	Missing	0001
AA2			1	English	0000
AA2			2	French	1000
AA2			3	Other - specify	0100
AA2			6	Valid skip	0010
B_D01a3DE1	13	Education National - Highest Level of Education -	-1	Missing	000000000001
B_D01a3DE1			1	No formal qualificat	000000000000
B_D01a3DE1			2	Left school in Germa	100000000000
B_D01a3DE1			3	German General educa	010000000000
B_D01a3DE1			4	German Evening schoo	001000000000
B_D01a3DE1			5	German vocational ed	000100000000
B_D01a3DE1			6	German university ed	000010000000
B_D01a3DE1			7	German other degree	000001000000
B_D01a3DE1			8	Left school in other	000000100000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_D01a3DE1			9	Foreign general educ	000000010000
B_D01a3DE1			10	Foreign vocational e	000000001000
B_D01a3DE1			11	Foreign university	000000000100
B_D01a3DE1			96	Valid skip	000000000010
B_D02b3DE1	7	Education National - Current Level of Education -	-1	Missing	000001
B_D02b3DE1			1	General education	000000
B_D02b3DE1			2	Evening school	100000
B_D02b3DE1			3	Vocational education	010000
B_D02b3DE1			4	University education	001000
B_D02b3DE1			5	German other degree	000100
B_D02b3DE1			6	Valid skip	000010
B_D03b3DE1	7	Education National - Uncompleted Education - Deriv	-1	Missing	000001
B_D03b3DE1			1	General education	000000
B_D03b3DE1			2	Evening school	100000
B_D03b3DE1			3	Vocational education	010000
B_D03b3DE1			4	University education	001000
B_D03b3DE1			5	German other degree	000100
B_D03b3DE1			6	Valid skip	000010
B_D05a3DE1	7	Education National - Formal Level of Education - D	-1	Missing	000001
B_D05a3DE1			1	General education	000000
B_D05a3DE1			2	Evening school	100000
B_D05a3DE1			3	Vocational education	010000
B_D05a3DE1			4	University education	001000
B_D05a3DE1			5	German other degree	000100
B_D05a3DE1			6	Valid skip	000010
B_Q00CZ01	4	Education - Level 01	-1	Missing	001
B_Q00CZ01			1	Yes	000
B_Q00CZ01			2	No	100
B_Q00CZ01			6	Valid skip	010
B_Q00CZ02	4	Education - Level 02	-1	Missing	001
B_Q00CZ02			1	Yes	000
B_Q00CZ02			2	No	100
B_Q00CZ02			6	Valid skip	010
B_Q00CZ03	4	Education - Level 03	-1	Missing	001
B_Q00CZ03			1	Yes	000
B_Q00CZ03			2	No	100
B_Q00CZ03			6	Valid skip	010
B_Q00CZ04	4	Education - Level 04	-1	Missing	001
B_Q00CZ04			1	Yes	000
B_Q00CZ04			2	No	100
B_Q00CZ04			6	Valid skip	010
B_Q00CZ05	4	Education - Level 05	-1	Missing	001
B_Q00CZ05			1	Yes	000
B_Q00CZ05			2	No	100
B_Q00CZ05			6	Valid skip	010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q00CZ06	4	Education - Level 06	-1	Missing	001
B_Q00CZ06			1	Yes	000
B_Q00CZ06			2	No	100
B_Q00CZ06			6	Valid skip	010
B_Q00CZ07	4	Education - Level 07	-1	Missing	001
B_Q00CZ07			1	Yes	000
B_Q00CZ07			2	No	100
B_Q00CZ07			6	Valid skip	010
B_Q00CZ08	4	Education - Level 08	-1	Missing	001
B_Q00CZ08			1	Yes	000
B_Q00CZ08			2	No	100
B_Q00CZ08			6	Valid skip	010
B_Q00CZ09	4	Education - Level 09	-1	Missing	001
B_Q00CZ09			1	Yes	000
B_Q00CZ09			2	No	100
B_Q00CZ09			6	Valid skip	010
B_Q00CZ10	4	Education - Level 10	-1	Missing	001
B_Q00CZ10			1	Yes	000
B_Q00CZ10			2	No	100
B_Q00CZ10			6	Valid skip	010
B_Q00CZ11	4	Education - Level 11	-1	Missing	001
B_Q00CZ11			1	Yes	000
B_Q00CZ11			2	No	100
B_Q00CZ11			6	Valid skip	010
B_Q00CZ12	4	Education - Level 12	-1	Missing	001
B_Q00CZ12			1	Yes	000
B_Q00CZ12			2	No	100
B_Q00CZ12			6	Valid skip	010
B_Q00CZ13	4	Education - Level 13	-1	Missing	001
B_Q00CZ13			1	Yes	000
B_Q00CZ13			2	No	100
B_Q00CZ13			6	Valid skip	010
B_Q00CZ14	4	Education - Level 14	-1	Missing	001
B_Q00CZ14			1	Yes	000
B_Q00CZ14			2	No	100
B_Q00CZ14			6	Valid skip	010
B_Q00SEX	4	Verification education	-1	Missing	001
B_Q00SEX			1	Yes	000
B_Q00SEX			2	No	100
B_Q00SEX			6	Valid skip	010
B_Q00UKX_01	4	Education - All qualifications - Degree level	-1	Missing	001
B_Q00UKX_01			1	Marked	000
B_Q00UKX_01			2	Not marked	100
B_Q00UKX_01			6	Valid skip	010
B_Q00UKX_02	4	Education - All qualifications - Diploma in HE	-1	Missing	001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q00UKX_02			1	Marked	000
B_Q00UKX_02			2	Not marked	100
B_Q00UKX_02			6	Valid skip	010
B_Q00UKX_03	4	Education - All qualifications - HNC/HND	-1	Missing	001
B_Q00UKX_03			1	Marked	000
B_Q00UKX_03			2	Not marked	100
B_Q00UKX_03			6	Valid skip	010
B_Q00UKX_04	4	Education - All qualifications - ONC/OND	-1	Missing	001
B_Q00UKX_04			1	Marked	000
B_Q00UKX_04			2	Not marked	100
B_Q00UKX_04			6	Valid skip	010
B_Q00UKX_05	4	Education - All qualifications - BTEC,BEC,TEC, EdE	-1	Missing	001
B_Q00UKX_05			1	Marked	000
B_Q00UKX_05			2	Not marked	100
B_Q00UKX_05			6	Valid skip	010
B_Q00UKX_06	4	Education - All qualifications - SCOTVEC,SCOTEC,SC	-1	Missing	001
B_Q00UKX_06			1	Marked	000
B_Q00UKX_06			2	Not marked	100
B_Q00UKX_06			6	Valid skip	010
B_Q00UKX_07	4	Education - All qualifications - Teaching qual exc	-1	Missing	001
B_Q00UKX_07			1	Marked	000
B_Q00UKX_07			2	Not marked	100
B_Q00UKX_07			6	Valid skip	010
B_Q00UKX_08	4	Education - All qualifications - Nursing or other	-1	Missing	001
B_Q00UKX_08			1	Marked	000
B_Q00UKX_08			2	Not marked	100
B_Q00UKX_08			6	Valid skip	010
B_Q00UKX_09	4	Education - All qualifications - Other HE qual bel	-1	Missing	001
B_Q00UKX_09			1	Marked	000
B_Q00UKX_09			2	Not marked	100
B_Q00UKX_09			6	Valid skip	010
B_Q00UKX_10	4	Education - All qualifications - A Level/vocationa	-1	Missing	001
B_Q00UKX_10			1	Marked	000
B_Q00UKX_10			2	Not marked	100
B_Q00UKX_10			6	Valid skip	010
B_Q00UKX_11	4	Education - All qualifications - highers (Scotland	-1	Missing	001
B_Q00UKX_11			1	Marked	000
B_Q00UKX_11			2	Not marked	100
B_Q00UKX_11			6	Valid skip	010
B_Q00UKX_12	4	Education - All qualifications - NVQ/SVQ	-1	Missing	001
B_Q00UKX_12			1	Marked	000
B_Q00UKX_12			2	Not marked	100
B_Q00UKX_12			6	Valid skip	010
B_Q00UKX_13	4	Education - All qualifications - GNVQ/GSVQ	-1	Missing	001
B_Q00UKX_13			1	Marked	000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q00UKX_13			2	Not marked	100
B_Q00UKX_13			6	Valid skip	010
B_Q00UKX_14	4	Education - All qualifications - AS Level/Vocation	-1	Missing	001
B_Q00UKX_14			1	Marked	000
B_Q00UKX_14			2	Not marked	100
B_Q00UKX_14			6	Valid skip	010
B_Q00UKX_15	4	Education - All qualifications - Advanced Highers/	-1	Missing	001
B_Q00UKX_15			1	Marked	000
B_Q00UKX_15			2	Not marked	100
B_Q00UKX_15			6	Valid skip	010
B_Q00UKX_16	4	Education - All qualifications - Access to HE	-1	Missing	001
B_Q00UKX_16			1	Marked	000
B_Q00UKX_16			2	Not marked	100
B_Q00UKX_16			6	Valid skip	010
B_Q00UKX_17	4	Education - All qualifications - O Level/GCSE/Voca	-1	Missing	001
B_Q00UKX_17			1	Marked	000
B_Q00UKX_17			2	Not marked	100
B_Q00UKX_17			6	Valid skip	010
B_Q00UKX_18	4	Education - All qualifications - Intermediate 1 or	-1	Missing	001
B_Q00UKX_18			1	Marked	000
B_Q00UKX_18			2	Not marked	100
B_Q00UKX_18			6	Valid skip	010
B_Q00UKX_19	4	Education - All qualifications - Standard Grade or	-1	Missing	001
B_Q00UKX_19			1	Marked	000
B_Q00UKX_19			2	Not marked	100
B_Q00UKX_19			6	Valid skip	010
B_Q00UKX_20	4	Education - All qualifications - National Qualific	-1	Missing	001
B_Q00UKX_20			1	Marked	000
B_Q00UKX_20			2	Not marked	100
B_Q00UKX_20			6	Valid skip	010
B_Q00UKX_21	4	Education - All qualifications - RSA/OCR	-1	Missing	001
B_Q00UKX_21			1	Marked	000
B_Q00UKX_21			2	Not marked	100
B_Q00UKX_21			6	Valid skip	010
B_Q00UKX_22	4	Education - All qualifications - City and Guilds	-1	Missing	001
B_Q00UKX_22			1	Marked	000
B_Q00UKX_22			2	Not marked	100
B_Q00UKX_22			6	Valid skip	010
B_Q00UKX_23	4	Education - All qualifications - YT Certificate/YT	-1	Missing	001
B_Q00UKX_23			1	Marked	000
B_Q00UKX_23			2	Not marked	100
B_Q00UKX_23			6	Valid skip	010
B_Q00UKX_24	4	Education - All qualifications - Key Skills/Basic	-1	Missing	001
B_Q00UKX_24			1	Marked	000
B_Q00UKX_24			2	Not marked	100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q00UKX_24			6	Valid skip	010
B_Q00UKX_25	4	Education - All qualifications - Entry Level quali	-1	Missing	001
B_Q00UKX_25			1	Marked	000
B_Q00UKX_25			2	Not marked	100
B_Q00UKX_25			6	Valid skip	010
B_Q00UKX_26	4	Education - All qualifications - Foreign Qualifica	-1	Missing	001
B_Q00UKX_26			1	Marked	000
B_Q00UKX_26			2	Not marked	100
B_Q00UKX_26			6	Valid skip	010
B_Q00UKX_27	4	Education - All qualifications - Any other profess	-1	Missing	001
B_Q00UKX_27			1	Marked	000
B_Q00UKX_27			2	Not marked	100
B_Q00UKX_27			6	Valid skip	010
B_Q00UKX_28	4	Education - All qualifications - No formal qualifi	-1	Missing	001
B_Q00UKX_28			1	Marked	000
B_Q00UKX_28			2	Not marked	100
B_Q00UKX_28			6	Valid skip	010
B_Q01a1AU	13	Education - Highest primary/secondary school - Cou	-1	Missing	000000000001
B_Q01a1AU			1	Australia	000000000000
B_Q01a1AU			2	England	100000000000
B_Q01a1AU			3	New Zealand	010000000000
B_Q01a1AU			4	Italy	001000000000
B_Q01a1AU			5	Viet Nam	000100000000
B_Q01a1AU			6	Scotland	000010000000
B_Q01a1AU			7	Greece	000001000000
B_Q01a1AU			8	Germany	000000100000
B_Q01a1AU			9	Philippines	000000010000
B_Q01a1AU			10	India	000000001000
B_Q01a1AU			11	Other - please speci	000000000100
B_Q01a1AU			96	Valid skip	000000000010
B_Q01a1AU12	14	Education - Highest primary/secondary - Month stop	-1	Missing	0000000000001
B_Q01a1AU12			1	January	0000000000000
B_Q01a1AU12			2	February	1000000000000
B_Q01a1AU12			3	March	0100000000000
B_Q01a1AU12			4	April	0010000000000
B_Q01a1AU12			5	May	0001000000000
B_Q01a1AU12			6	June	0000100000000
B_Q01a1AU12			7	July	0000010000000
B_Q01a1AU12			8	August	0000001000000
B_Q01a1AU12			9	September	0000000100000
B_Q01a1AU12			10	October	0000000010000
B_Q01a1AU12			11	November	0000000001000
B_Q01a1AU12			12	December	0000000000100
B_Q01a1AU12			96	Valid skip	0000000000010
B_Q01a1AU7	4	Education - Ever started but did not complete a le	-1	Missing	001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a1AU7			1	Yes	000
B_Q01a1AU7			2	No	100
B_Q01a1AU7			6	Valid skip	010
B_Q01a1AU8	8	Education - Highest primary/secondary school - Stu	-1	Missing	0000001
B_Q01a1AU8			1	Year 12 or equivalen	0000000
B_Q01a1AU8			2	Year 11 or equivalen	1000000
B_Q01a1AU8			3	Year 10 or equivalen	0100000
B_Q01a1AU8			4	Year 9 or equivalent	0010000
B_Q01a1AU8			5	Year 8 or equivalent	0001000
B_Q01a1AU8			6	Year 7 or below	0000100
B_Q01a1AU8			96	Valid skip	0000010
B_Q01a2AT	8	Education - Highest qualification - Country of for	-1	Missing	0000001
B_Q01a2AT			1	Turkey	0000000
B_Q01a2AT			2	Serbia	1000000
B_Q01a2AT			3	Bosnia-Herzogovina	0100000
B_Q01a2AT			4	Croatia	0010000
B_Q01a2AT			5	Germany	0001000
B_Q01a2AT			6	Other country	0000100
B_Q01a2AT			96	Valid skip	0000010
B_Q01a2AU	12	Education - Highest qualification - Country comple	-1	Missing	00000000001
B_Q01a2AU			1	England	00000000000
B_Q01a2AU			2	New Zealand	10000000000
B_Q01a2AU			3	Italy	01000000000
B_Q01a2AU			4	Viet Nam	00100000000
B_Q01a2AU			5	Scotland	00010000000
B_Q01a2AU			6	Greece	00001000000
B_Q01a2AU			7	Germany	00000100000
B_Q01a2AU			8	Philippines	00000010000
B_Q01a2AU			9	India	00000001000
B_Q01a2AU			10	Other - please speci	00000000100
B_Q01a2AU			96	Valid skip	00000000010
B_Q01a2BE	12	Education - Highest qualification - Country of for	-1	Missing	00000000001
B_Q01a2BE			1	The Netherlands	00000000000
B_Q01a2BE			2	Italy	10000000000
B_Q01a2BE			3	France	01000000000
B_Q01a2BE			4	Germany	00100000000
B_Q01a2BE			5	Spain	00010000000
B_Q01a2BE			6	Morocco	00001000000
B_Q01a2BE			7	Turkey	00000100000
B_Q01a2BE			8	Poland	00000010000
B_Q01a2BE			9	Former Yugoslavia	00000001000
B_Q01a2BE			10	Other country	00000000100
B_Q01a2BE			96	Valid skip	00000000010
B_Q01a2CY	9	Education - Highest qualification - Country of for	-1	Missing	00000001
B_Q01a2CY			1	Cyprus	00000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a2CY			2	Greece	10000000
B_Q01a2CY			3	United Kingdom	01000000
B_Q01a2CY			4	USA	00100000
B_Q01a2CY			5	Russian Federation	00010000
B_Q01a2CY			6	France	00001000
B_Q01a2CY			7	Other Country	00000100
B_Q01a2CY			96	Valid skip	00000010
B_Q01a2CZ	9	Education - Highest qualification - Country of for	-1	Missing	00000001
B_Q01a2CZ			1	Country 1	00000000
B_Q01a2CZ			2	Country 2	10000000
B_Q01a2CZ			3	Country 3	01000000
B_Q01a2CZ			4	Country 4	00100000
B_Q01a2CZ			5	Country 5	00010000
B_Q01a2CZ			6	Country 6	00001000
B_Q01a2CZ			7	Other country	00000100
B_Q01a2CZ			96	Valid skip	00000010
B_Q01a2DE2	13	Education National - Highest qualification - Count	-1	Missing	000000000001
B_Q01a2DE2			1	Turkey	000000000000
B_Q01a2DE2			2	Italy	100000000000
B_Q01a2DE2			3	Poland	010000000000
B_Q01a2DE2			4	Greece	001000000000
B_Q01a2DE2			5	Serbia	000100000000
B_Q01a2DE2			6	Croatia	000010000000
B_Q01a2DE2			7	Russian Federation	000001000000
B_Q01a2DE2			8	Bosnia and Herzegovi	000000100000
B_Q01a2DE2			9	United Kingdom	000000010000
B_Q01a2DE2			10	United States	000000001000
B_Q01a2DE2			11	Another country	000000000100
B_Q01a2DE2			96	Valid skip	000000000010
B_Q01a2DK	9	In which country did you gain this qualification?	-1	Missing	00000001
B_Q01a2DK			1	Turkey	00000000
B_Q01a2DK			2	Germany	10000000
B_Q01a2DK			3	Poland	01000000
B_Q01a2DK			4	Iraq	00100000
B_Q01a2DK			5	Bosnia-Herzegovinia	00010000
B_Q01a2DK			6	Norway	00001000
B_Q01a2DK			7	Other country	00000100
B_Q01a2DK			96	Valid skip	00000010
B_Q01a2EE	9	Education - Highest qualification - Country of for	-1	Missing	00000001
B_Q01a2EE			1	Russia	00000000
B_Q01a2EE			2	USA	10000000
B_Q01a2EE			3	Germany	01000000
B_Q01a2EE			4	UK	00100000
B_Q01a2EE			5	Finland	00010000
B_Q01a2EE			6	Sweden	00001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a2EE			7	Other country	00000100
B_Q01a2EE			96	Valid skip	00000010
B_Q01a2ES	13	Education - Highest qualification - Country of for	-1	Missing	000000000001
B_Q01a2ES			1	Argentina	000000000000
B_Q01a2ES			2	Colombia	100000000000
B_Q01a2ES			3	Ecuador	010000000000
B_Q01a2ES			4	Marruecos	001000000000
B_Q01a2ES			5	Marruecos	000100000000
B_Q01a2ES			6	Marrblica Dominicana	000010000000
B_Q01a2ES			7	Marrbla	000001000000
B_Q01a2ES			8	Venezuela	000000100000
B_Q01a2ES			9	Reino Unido	000000010000
B_Q01a2ES			10	Alemania	000000001000
B_Q01a2ES			11	Alemanias	000000000100
B_Q01a2ES			96	Valid skip	000000000010
B_Q01a2FI	7	Education - Highest qualification - Country of for	-1	Missing	000001
B_Q01a2FI			1	Sweden	000000
B_Q01a2FI			2	Russia	100000
B_Q01a2FI			3	Former Soviet Union	010000
B_Q01a2FI			4	Estonia	001000
B_Q01a2FI			5	Other country	000100
B_Q01a2FI			96	Valid skip	000010
B_Q01a2FR	12	Education - Highest qualification - Country of for	-1	Missing	00000000001
B_Q01a2FR			1	Algeria	00000000000
B_Q01a2FR			2	Germany	10000000000
B_Q01a2FR			3	Spain	01000000000
B_Q01a2FR			4	Italy	00100000000
B_Q01a2FR			5	Morocco	00010000000
B_Q01a2FR			6	Portugal	00001000000
B_Q01a2FR			7	United Kingdom	00000100000
B_Q01a2FR			8	Tunisia	00000010000
B_Q01a2FR			9	Turkey	00000001000
B_Q01a2FR			10	Other countries	00000000100
B_Q01a2FR			96	Valid skip	00000000010
B_Q01a2IE	10	Education - Highest qualification - Country of for	-1	Missing	000000001
B_Q01a2IE			1	Poland	000000000
B_Q01a2IE			2	United Kingdom	100000000
B_Q01a2IE			3	Lithuania	010000000
B_Q01a2IE			4	Latvia	001000000
B_Q01a2IE			5	Germany	000100000
B_Q01a2IE			6	Romania	000010000
B_Q01a2IE			7	Northern Ireland	000001000
B_Q01a2IE			8	Other country	000000100
B_Q01a2IE			96	Valid skip	000000010
B_Q01a2IT	18	Education - Highest qualification - Country of for	-1	Missing	00000000000000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a2IT			1	Albania	0000000000000000
B_Q01a2IT			2	China	1000000000000000
B_Q01a2IT			3	Ecuador	0100000000000000
B_Q01a2IT			4	Philippines	0010000000000000
B_Q01a2IT			5	France	0001000000000000
B_Q01a2IT			6	Germany	0000100000000000
B_Q01a2IT			7	Morocco	0000010000000000
B_Q01a2IT			8	Peru	0000001000000000
B_Q01a2IT			9	Poland	0000000100000000
B_Q01a2IT			10	United Kingdom	0000000010000000
B_Q01a2IT			11	Romania	0000000001000000
B_Q01a2IT			12	Spain	0000000000100000
B_Q01a2IT			13	United States of Ame	0000000000010000
B_Q01a2IT			14	Tunisia	0000000000001000
B_Q01a2IT			15	Ukraina	0000000000000100
B_Q01a2IT			16	Other	0000000000000010
B_Q01a2IT			96	Valid skip	0000000000000001
B_Q01a2JP	12	Education - Highest qualification - Country of for	-1	Missing	00000000001
B_Q01a2JP			1	USA	00000000000
B_Q01a2JP			2	Canada	10000000000
B_Q01a2JP			3	UK	01000000000
B_Q01a2JP			4	Australia	00100000000
B_Q01a2JP			5	New Zealand	00010000000
B_Q01a2JP			6	Republic of Korea	00001000000
B_Q01a2JP			7	China	00000100000
B_Q01a2JP			8	Germany	00000010000
B_Q01a2JP			9	France	00000001000
B_Q01a2JP			10	Other country	00000000100
B_Q01a2JP			96	Valid skip	00000000010
B_Q01a2KO	9	KO_Education - earned country	-1	Missing	00000001
B_Q01a2KO			1	China	00000000
B_Q01a2KO			2	United States	10000000
B_Q01a2KO			3	Vietnam	01000000
B_Q01a2KO			4	Philippines	00100000
B_Q01a2KO			5	Thailand	00010000
B_Q01a2KO			6	Japan	00001000
B_Q01a2KO			7	Other country	00000100
B_Q01a2KO			96	Valid skip	00000010
B_Q01a2NL	9	Education - Highest qualification - Country of for	-1	Missing	00000001
B_Q01a2NL			1	Marocco	00000000
B_Q01a2NL			2	Turkey	10000000
B_Q01a2NL			3	Germany	01000000
B_Q01a2NL			4	Belgium	00100000
B_Q01a2NL			5	France	00010000
B_Q01a2NL			6	United Kingdom	00001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a2NL			7	Other country	00000100
B_Q01a2NL			96	Valid skip	00000010
B_Q01a2NO	9	Education - Highest qualification - Country of for	-1	Missing	00000001
B_Q01a2NO			1	Australia	00000000
B_Q01a2NO			2	Denmark	10000000
B_Q01a2NO			3	Pakistan	01000000
B_Q01a2NO			4	UK	00100000
B_Q01a2NO			5	Sweden	00010000
B_Q01a2NO			6	USA	00001000
B_Q01a2NO			7	Other country	00000100
B_Q01a2NO			96	Valid skip	00000010
B_Q01a2PL	15	Education - Highest qualification - Country of for	-1	Missing	00000000000001
B_Q01a2PL			1	Belarus	00000000000000
B_Q01a2PL			2	Czech Republic	10000000000000
B_Q01a2PL			3	England	01000000000000
B_Q01a2PL			4	France	00100000000000
B_Q01a2PL			5	Germany	00010000000000
B_Q01a2PL			6	Lithuania	00001000000000
B_Q01a2PL			7	Netherlands	00000100000000
B_Q01a2PL			8	Poland	00000010000000
B_Q01a2PL			9	Russia	00000001000000
B_Q01a2PL			10	Slovakia	00000000100000
B_Q01a2PL			11	Ukraine	00000000010000
B_Q01a2PL			12	United States of Ame	00000000001000
B_Q01a2PL			13	Other country	00000000000100
B_Q01a2PL			96	Valid skip	00000000000010
B_Q01a2RU	9	Education - Highest qualification - Country of for	-1	Missing	00000001
B_Q01a2RU			1	Country 1	00000000
B_Q01a2RU			2	Country 2	10000000
B_Q01a2RU			3	Country 3	01000000
B_Q01a2RU			4	Country 4	00100000
B_Q01a2RU			5	Country 5	00010000
B_Q01a2RU			6	Country 6	00001000
B_Q01a2RU			7	Other country	00000100
B_Q01a2RU			96	Valid skip	00000010
B_Q01a2SE	15	Education - Highest qualification - Country of for	-1	Missing	00000000000001
B_Q01a2SE			1	Finland	00000000000000
B_Q01a2SE			2	Irak	10000000000000
B_Q01a2SE			3	Serbien	01000000000000
B_Q01a2SE			4	Iran	00100000000000
B_Q01a2SE			5	Polen	00010000000000
B_Q01a2SE			6	Bosnien-Hercegovina	00001000000000
B_Q01a2SE			7	Turkiet	00000100000000
B_Q01a2SE			8	Danmark	00000010000000
B_Q01a2SE			9	Norge	00000001000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a2SE			10	Chile	00000000100000
B_Q01a2SE			11	Tyskland	00000000010000
B_Q01a2SE			12	Kroatien	00000000001000
B_Q01a2SE			13	Annat land var god a	00000000000100
B_Q01a2SE			96	Valid skip	00000000000010
B_Q01a2SK	9	Education - Highest qualification - Country of for	-1	Missing	00000001
B_Q01a2SK			1	Czech republic	00000000
B_Q01a2SK			2	Hungary	10000000
B_Q01a2SK			3	Austria	01000000
B_Q01a2SK			4	Poland	00100000
B_Q01a2SK			5	Russia	00010000
B_Q01a2SK			6	Great Britain	00001000
B_Q01a2SK			7	other country	00000100
B_Q01a2SK			96	Valid skip	00000010
B_Q01a2UK	15	Education - Highest qualification - Country of for	-1	Missing	00000000000001
B_Q01a2UK			1	India	00000000000000
B_Q01a2UK			2	Poland	10000000000000
B_Q01a2UK			3	Pakistan	01000000000000
B_Q01a2UK			4	Germany	00100000000000
B_Q01a2UK			5	South Africa	00010000000000
B_Q01a2UK			6	Bangladesh	00001000000000
B_Q01a2UK			7	Nigeria	00000100000000
B_Q01a2UK			8	Kenya	00000010000000
B_Q01a2UK			9	United States	00000001000000
B_Q01a2UK			10	Phillippines	00000000100000
B_Q01a2UK			11	France	00000000010000
B_Q01a2UK			12	Australia	00000000001000
B_Q01a2UK			13	Other Country	00000000000100
B_Q01a2UK			96	Valid skip	00000000000010
B_Q01a2US	9	Education - Highest qualification - Country of for	-1	Missing	00000001
B_Q01a2US			1	Mexico	00000000
B_Q01a2US			2	China	10000000
B_Q01a2US			3	Phillipines	01000000
B_Q01a2US			4	India	00100000
B_Q01a2US			5	Russia	00010000
B_Q01a2US			6	Colombia	00001000
B_Q01a2US			7	Other country	00000100
B_Q01a2US			96	Valid skip	00000010
B_Q01a3AT	18	Education - Highest qualification - Level of forei	-1	Missing	00000000000000001
B_Q01a3AT			1	No compulsory school	00000000000000000
B_Q01a3AT			2	Compulsory school	10000000000000000
B_Q01a3AT			3	Apprenticeship	01000000000000000
B_Q01a3AT			4	Vocational School (<	00100000000000000
B_Q01a3AT			5	Vocational School (2	00010000000000000
B_Q01a3AT			6	Nursing	00001000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a3AT			7	Master craftsman's c	00000100000000000
B_Q01a3AT			8	Academic Secondary S	00000010000000000
B_Q01a3AT			9	Vocational college	00000001000000000
B_Q01a3AT			10	Post-secondary cours	00000000100000000
B_Q01a3AT			11	Post-secondary colle	00000000010000000
B_Q01a3AT			12	University courses	00000000001000000
B_Q01a3AT			13	University-Bachelor	00000000000100000
B_Q01a3AT			14	University-Master	00000000000010000
B_Q01a3AT			15	Post-graduate course	00000000000001000
B_Q01a3AT			16	Doctoral Programme	00000000000000100
B_Q01a3AT			96	Valid skip	00000000000000010
B_Q01a3AU	12	Education - Highest qualification - Level of quali	-1	Missing	00000000001
B_Q01a3AU			1	Certificate I	00000000000
B_Q01a3AU			2	Certificate II	10000000000
B_Q01a3AU			3	Certificate III	01000000000
B_Q01a3AU			4	Certificate IV	00100000000
B_Q01a3AU			5	Diploma	00010000000
B_Q01a3AU			6	Advanced Diploma and	00001000000
B_Q01a3AU			7	Bachelor degree (inc	00000100000
B_Q01a3AU			8	Graduate Diploma or	00000010000
B_Q01a3AU			9	Masters	00000001000
B_Q01a3AU			10	Doctorate	00000000100
B_Q01a3AU			96	Valid skip	00000000010
B_Q01a3BE	13	Education - Highest qualification - Level of forei	-1	Missing	000000000001
B_Q01a3BE			1	No formal qualificat	000000000000
B_Q01a3BE			2	ISCED 1	100000000000
B_Q01a3BE			3	ISCED 2	010000000000
B_Q01a3BE			4	ISCED 3C 2 years or	001000000000
B_Q01a3BE			5	ISCED 3A-B	000100000000
B_Q01a3BE			6	ISCED 3 (without dis	000010000000
B_Q01a3BE			7	ISCED 4A-B	000001000000
B_Q01a3BE			8	ISCED 5B	000000100000
B_Q01a3BE			9	ISCED 5A, bachelor d	000000010000
B_Q01a3BE			10	ISCED 5A, master deg	000000001000
B_Q01a3BE			11	ISCED 6	000000000100
B_Q01a3BE			96	Valid skip	000000000010
B_Q01a3CY	10	Education - Highest qualification - Level of forei	-1	Missing	000000001
B_Q01a3CY			1	I never went to scho	000000000
B_Q01a3CY			2	Primary school	100000000
B_Q01a3CY			3	Public/Private Secon	010000000
B_Q01a3CY			4	High School/Vocation	001000000
B_Q01a3CY			5	Non-Univ. Degree/Dip	000100000
B_Q01a3CY			6	Undergraduate degree	000010000
B_Q01a3CY			7	Postgraduate degree,	000001000
B_Q01a3CY			8	Doctorate	000000100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a3CY			96	Valid skip	000000010
B_Q01a3CZ	15	Education - Highest qualification - Level of forei	-1	Missing	00000000000001
B_Q01a3CZ			1	No formal education	00000000000000
B_Q01a3CZ			2	First level of basic	10000000000000
B_Q01a3CZ			3	basic ISCED 2	01000000000000
B_Q01a3CZ			4	vocational without m	00100000000000
B_Q01a3CZ			5	vocational without m	00010000000000
B_Q01a3CZ			6	ISCED 3A vocational	00001000000000
B_Q01a3CZ			7	ISCED 3A technical w	00000100000000
B_Q01a3CZ			8	ISCED 3A general wit	00000010000000
B_Q01a3CZ			9	ISCED 4 follow-up co	00000001000000
B_Q01a3CZ			10	ISCED 5B higher prof	00000000100000
B_Q01a3CZ			11	ISCED 5A, bachelor	00000000010000
B_Q01a3CZ			12	ISCED 5A, master	00000000001000
B_Q01a3CZ			13	ISCED 6, post gradua	00000000000100
B_Q01a3CZ			96	Valid skip	00000000000010
B_Q01a3DE2a	8	Education National - Highest school qualification	-1	Missing	0000001
B_Q01a3DE2a			1	Left school without	0000000
B_Q01a3DE2a			2	Hauptschulabschluss	1000000
B_Q01a3DE2a			3	Realschulabschluss (	0100000
B_Q01a3DE2a			4	Fachhochschulreife,	0010000
B_Q01a3DE2a			5	Abitur/EOS (General	0001000
B_Q01a3DE2a			6	Did not attend schoo	0000100
B_Q01a3DE2a			96	Valid skip	0000010
B_Q01a3DE2b	11	Education National - Highest professional qualific	-1	Missing	0000000001
B_Q01a3DE2b			1	Apprenticeship (Lehr	0000000000
B_Q01a3DE2b			2	Basic vocational tra	1000000000
B_Q01a3DE2b			3	Training at Fachschu	0100000000
B_Q01a3DE2b			4	Berufsakademie, Fach	0010000000
B_Q01a3DE2b			5	Bachelor at Fachhoch	0001000000
B_Q01a3DE2b			6	Master/Diplom at Fac	0000100000
B_Q01a3DE2b			7	Bachelor at universi	0000010000
B_Q01a3DE2b			8	Master/Diplom at uni	0000001000
B_Q01a3DE2b			9	Doctorate	0000000100
B_Q01a3DE2b			96	Valid skip	0000000010
B_Q01a3DK	16	Can you indicate which level in our national educa	-1	Missing	000000000000001
B_Q01a3DK			1	No formal education	000000000000000
B_Q01a3DK			2	Primary school, grad	100000000000000
B_Q01a3DK			3	Lower secondary, gra	010000000000000
B_Q01a3DK			4	Upper secondary voca	001000000000000
B_Q01a3DK			5	Upper secondary voca	000100000000000
B_Q01a3DK			6	Upper secondary gene	000010000000000
B_Q01a3DK			7	Upper secondary unde	000001000000000
B_Q01a3DK			8	Post secondary short	000000100000000
B_Q01a3DK			9	Post secondary entra	000000010000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a3DK			10	Post secondary non t	000000001000000
B_Q01a3DK			11	Tertiary not researc	000000000100000
B_Q01a3DK			12	Bachelor degree	000000000010000
B_Q01a3DK			13	Master degree	000000000001000
B_Q01a3DK			14	Ph.d or otther resea	000000000000100
B_Q01a3DK			96	Valid skip	000000000000010
B_Q01a3EE	20	Education - Highest qualification - Level of forei	-1	Missing	0000000000000000001
B_Q01a3EE			1	Without primary educ	0000000000000000000
B_Q01a3EE			2	Primary education	1000000000000000000
B_Q01a3EE			3	Basic education	0100000000000000000
B_Q01a3EE			4	General secondary ed	0010000000000000000
B_Q01a3EE			5	Vocational education	0001000000000000000
B_Q01a3EE			6	Vocational education	0000100000000000000
B_Q01a3EE			7	Vocational education	0000010000000000000
B_Q01a3EE			8	Vocational secondary	0000001000000000000
B_Q01a3EE			9	Secondary specialise	0000000100000000000
B_Q01a3EE			10	Vocational secondary	0000000010000000000
B_Q01a3EE			11	Secondary specialise	0000000001000000000
B_Q01a3EE			12	Applied higher educa	0000000000100000000
B_Q01a3EE			13	Bachelor's degree (3	0000000000010000000
B_Q01a3EE			14	Bachelor's degree (4	0000000000001000000
B_Q01a3EE			15	Higher education (st	0000000000000100000
B_Q01a3EE			16	Master's degree (3+2	0000000000000010000
B_Q01a3EE			17	Master's degree (4+2	0000000000000001000
B_Q01a3EE			18	Doctoral degree (inc	0000000000000000100
B_Q01a3EE			96	Valid skip	0000000000000000010
B_Q01a3FI	13	Education - Highest qualification - Level of forei	-1	Missing	000000000001
B_Q01a3FI			1	No formal qualificat	000000000000
B_Q01a3FI			2	ISCED 1	100000000000
B_Q01a3FI			3	ISCED 2	010000000000
B_Q01a3FI			4	Upper secondary voca	001000000000
B_Q01a3FI			5	General upper second	000100000000
B_Q01a3FI			6	Specialist vocationa	000010000000
B_Q01a3FI			7	Vocational post-seco	000001000000
B_Q01a3FI			8	Polytechnic degree (	000000100000
B_Q01a3FI			9	Bachelor's degree (I	000000010000
B_Q01a3FI			10	Master's degree (ISC	000000001000
B_Q01a3FI			11	Licentiate's and doc	000000000100
B_Q01a3FI			96	Valid skip	000000000010
B_Q01a3FR	16	Education - Highest qualification - Level of forei	-1	Missing	000000000000001
B_Q01a3FR			1	No formal qualificat	000000000000000
B_Q01a3FR			2	ISCED 1	100000000000000
B_Q01a3FR			3	ISCED 2	010000000000000
B_Q01a3FR			4	ISCED 3C shorter tha	001000000000000
B_Q01a3FR			5	ISCED 3C 2 years or	000100000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a3FR			6	ISCED 3A-B	000010000000000
B_Q01a3FR			7	ISCED 3 (without dis	000001000000000
B_Q01a3FR			8	ISCED 4C	000000100000000
B_Q01a3FR			9	ISCED 4A-B	000000010000000
B_Q01a3FR			10	ISCED 4 (without dis	000000001000000
B_Q01a3FR			11	ISCED 5B	000000000100000
B_Q01a3FR			12	ISCED 5A, bachelor d	000000000010000
B_Q01a3FR			13	ISCED 5A, master deg	000000000001000
B_Q01a3FR			14	ISCED 6	000000000000100
B_Q01a3FR			96	Valid skip	000000000000010
B_Q01a3IE	15	Education - Highest qualification - Level of forei	-1	Missing	00000000000001
B_Q01a3IE			1	No formal education	00000000000000
B_Q01a3IE			2	Primary education (o	10000000000000
B_Q01a3IE			3	Secondary 1 (Junior/	01000000000000
B_Q01a3IE			4	Transition year prog	00100000000000
B_Q01a3IE			5	Secondary 2 (Leaving	00010000000000
B_Q01a3IE			6	Technical or Vocatio	00001000000000
B_Q01a3IE			7	Advanced Certificate	00000100000000
B_Q01a3IE			8	Higher Certificate (	00000010000000
B_Q01a3IE			9	Diploma (e.g. Nation	00000001000000
B_Q01a3IE			10	Honours Bachelor Deg	00000000100000
B_Q01a3IE			11	Professional (Honour	00000000010000
B_Q01a3IE			12	Post-Graduate (e.g.	00000000001000
B_Q01a3IE			13	Doctorate or higher	00000000000100
B_Q01a3IE			96	Valid skip	00000000000010
B_Q01a3IT	13	Education - Highest qualification - Level of forei	-1	Missing	000000000001
B_Q01a3IT			1	No formal qualificat	000000000000
B_Q01a3IT			2	Primary education or	100000000000
B_Q01a3IT			3	Lower secondary or s	010000000000
B_Q01a3IT			4	Regional Vocational	001000000000
B_Q01a3IT			5	Educational and voca	000100000000
B_Q01a3IT			6	Upper secondary educ	000010000000
B_Q01a3IT			7	Post-second. non ter	000001000000
B_Q01a3IT			8	Music Conservatory D	000000100000
B_Q01a3IT			9	First stage of terti	000000010000
B_Q01a3IT			10	First or second leve	000000001000
B_Q01a3IT			11	Research Doctoral de	000000000100
B_Q01a3IT			96	Valid skip	000000000010
B_Q01a3JP	16	Education - Highest qualification - Level of forei	-1	Missing	000000000000001
B_Q01a3JP			1	No formal school edu	000000000000000
B_Q01a3JP			2	Elementary school	100000000000000
B_Q01a3JP			3	Lower secondary scho	010000000000000
B_Q01a3JP			4	Short-term course of	001000000000000
B_Q01a3JP			5	Specialized course o	000100000000000
B_Q01a3JP			6	General/integrated c	000010000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a3JP			7	Passed upper seconda	000001000000000
B_Q01a3JP			8	Advanced course of u	000000100000000
B_Q01a3JP			9	Regular/advanced cou	000000010000000
B_Q01a3JP			10	Undergraducate progr	000000001000000
B_Q01a3JP			11	Master's programs/Do	000000000100000
B_Q01a3JP			12	Completed all work o	000000000010000
B_Q01a3JP			13	Doctoral programs of	000000000001000
B_Q01a3JP			14	Specialized training	000000000000100
B_Q01a3JP			96	Valid skip	000000000000010
B_Q01a3KO	13	KO_Education -comparision	-1	Missing	000000000001
B_Q01a3KO			1	no formal education	000000000000
B_Q01a3KO			2	Elementary school	100000000000
B_Q01a3KO			3	Middle school	010000000000
B_Q01a3KO			4	High school(college	001000000000
B_Q01a3KO			5	High school(vocation	000100000000
B_Q01a3KO			6	2-3 year college	000010000000
B_Q01a3KO			7	4 year college(speci	000001000000
B_Q01a3KO			8	4 year college(gener	000000100000
B_Q01a3KO			9	Master's degree(spec	000000010000
B_Q01a3KO			10	Master's degree(gene	000000001000
B_Q01a3KO			11	Doctoral degree	000000000100
B_Q01a3KO			96	Valid skip	000000000010
B_Q01a3NL	18	Education - Highest qualification - Level of forei	-1	Missing	00000000000000001
B_Q01a3NL			1	no formal qualificat	00000000000000000
B_Q01a3NL			2	primary education (i	10000000000000000
B_Q01a3NL			3	sec education, first	01000000000000000
B_Q01a3NL			4	sec education, first	00100000000000000
B_Q01a3NL			5	secondary education,	00010000000000000
B_Q01a3NL			6	secondary education,	00001000000000000
B_Q01a3NL			7	secondary education,	00000100000000000
B_Q01a3NL			8	secondary education,	00000010000000000
B_Q01a3NL			9	secondary education,	00000001000000000
B_Q01a3NL			10	secondary education,	00000000100000000
B_Q01a3NL			11	secondary education,	00000000010000000
B_Q01a3NL			12	tertiary education,	00000000001000000
B_Q01a3NL			13	tertiary education,	00000000000100000
B_Q01a3NL			14	tertiary education,	00000000000010000
B_Q01a3NL			15	tertiary education,	00000000000001000
B_Q01a3NL			16	tertiary education,	00000000000000100
B_Q01a3NL			96	Valid skip	00000000000000010
B_Q01a3NO	14	Education - Highest qualification - Level of forei	-1	Missing	0000000000001
B_Q01a3NO			1	No formal qualificat	0000000000000
B_Q01a3NO			2	ISCED 1	1000000000000
B_Q01a3NO			3	ISCED 2	0100000000000
B_Q01a3NO			4	ISCED 3C shorter tha	0010000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a3NO			5	ISCED 3C 2 years or	0001000000000
B_Q01a3NO			6	ISCED 3A-B	0000100000000
B_Q01a3NO			7	ISCED 4C	0000010000000
B_Q01a3NO			8	ISCED 4A-B	0000001000000
B_Q01a3NO			9	ISCED 5B	0000000100000
B_Q01a3NO			10	ISCED 5A, bachelor d	0000000010000
B_Q01a3NO			11	ISCED 5A, Master deg	0000000001000
B_Q01a3NO			12	ISCED 6	0000000000100
B_Q01a3NO			96	Valid skip	0000000000010
B_Q01a3PL	12	Education - Highest qualification - Level of forei	-1	Missing	00000000001
B_Q01a3PL			1	No formal qualificat	00000000000
B_Q01a3PL			2	ISCED 1	10000000000
B_Q01a3PL			3	ISCED 2	01000000000
B_Q01a3PL			4	ISCED 3C	00100000000
B_Q01a3PL			5	ISCED 3B	00010000000
B_Q01a3PL			6	ISCED 3A	00001000000
B_Q01a3PL			7	ISCED 4	00000100000
B_Q01a3PL			8	BA, ISCED 5A (I degr	00000010000
B_Q01a3PL			9	MA, ISCED 5A (II deg	00000001000
B_Q01a3PL			10	ISCED 6	00000000100
B_Q01a3PL			96	Valid skip	00000000010
B_Q01a3RU	11	Education - Highest qualification - Level of forei	-1	Missing	0000000001
B_Q01a3RU			1	No formal qualificat	0000000000
B_Q01a3RU			2	ISCED 1	1000000000
B_Q01a3RU			3	ISCED 2	0100000000
B_Q01a3RU			4	ISCED 3 (without dis	0010000000
B_Q01a3RU			5	ISCED 4 (without dis	0001000000
B_Q01a3RU			6	ISCED 5B	0000100000
B_Q01a3RU			7	ISCED 5A, bachelor d	0000010000
B_Q01a3RU			8	ISCED 5A, master deg	0000001000
B_Q01a3RU			9	ISCED 6	0000000100
B_Q01a3RU			96	Valid skip	0000000010
B_Q01a3SE1	18	Education correspondance	-1	Missing	00000000000000001
B_Q01a3SE1			1	Not stated or inferr	00000000000000000
B_Q01a3SE1			2	Not stated or inr	10000000000000000
B_Q01a3SE1			3	Grundskola, enhetssk	01000000000000000
B_Q01a3SE1			4	Yrkesutbildning	00100000000000000
B_Q01a3SE1			5	Grundskolekompetens	00010000000000000
B_Q01a3SE1			6	Flickskola	00001000000000000
B_Q01a3SE1			7	Gymnasie fackskola y	00000100000000000
B_Q01a3SE1			8	Gymnasie fackskola y	00000010000000000
B_Q01a3SE1			9	Gymnasie fackskola y	00000001000000000
B_Q01a3SE1			10	Vuxenutbildning mots	00000000100000000
B_Q01a3SE1			11	Vuxenutbildning mots	00000000010000000
B_Q01a3SE1			12	Eftergymnasial utbil	00000000001000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a3SE1			13	Eftergymnasial utbil	00000000000100000
B_Q01a3SE1			14	Eftergymnasial utbil	00000000000010000
B_Q01a3SE1			15	Eftergymnasial utbil	00000000000001000
B_Q01a3SE1			16	Forskarutbildning	00000000000000100
B_Q01a3SE1			96	Valid skip	00000000000000010
B_Q01a3SE2	4	Degree univ/coll	-1	Missing	001
B_Q01a3SE2			1	Yes	000
B_Q01a3SE2			2	No	100
B_Q01a3SE2			6	Valid skip	010
B_Q01a3SE3	6	Type of degree	-1	Missing	00001
B_Q01a3SE3			1	Fil Kand	00000
B_Q01a3SE3			2	Fil Mag	10000
B_Q01a3SE3			3	Master	01000
B_Q01a3SE3			4	Annan typ av examen	00100
B_Q01a3SE3			6	Valid skip	00010
B_Q01a3SK	13	Education - Highest qualification - Level of forei	-1	Missing	000000000001
B_Q01a3SK			1	Pre school education	000000000000
B_Q01a3SK			2	Primary school 1-4.	100000000000
B_Q01a3SK			3	Primary school 5.-9.	010000000000
B_Q01a3SK			4	Secondary technical	001000000000
B_Q01a3SK			5	Secondary technical	000100000000
B_Q01a3SK			6	Secondary schools wi	000010000000
B_Q01a3SK			7	Upper secondary scho	000001000000
B_Q01a3SK			8	Pre-tertiary school,	000000100000
B_Q01a3SK			9	Bachelor degree, Gra	000000010000
B_Q01a3SK			10	Master degree	000000001000
B_Q01a3SK			11	PhD studies, Second	000000000100
B_Q01a3SK			96	Valid skip	000000000010
B_Q01a3UK	11	Education - Highest qualification - Level of forei	-1	Missing	0000000001
B_Q01a3UK			1	No qualifications	0000000000
B_Q01a3UK			2	Key Skills, Basic sk	1000000000
B_Q01a3UK			3	O levels, GCSE or eq	0100000000
B_Q01a3UK			4	NVQ Level2, City & G	0010000000
B_Q01a3UK			5	A Levels or equivale	0001000000
B_Q01a3UK			6	Trade apprenticeship	0000100000
B_Q01a3UK			7	NVQ Level 3, City &	0000010000
B_Q01a3UK			8	Degree or higher deg	0000001000
B_Q01a3UK			9	NVQ Level 4 or 5, HN	0000000100
B_Q01a3UK			96	Valid skip	0000000010
B_Q01a3US	13	Education - Highest qualification - Level of forei	-1	Missing	000000000001
B_Q01a3US			1	Pre-primary or no sc	000000000000
B_Q01a3US			2	Grades 1-6	100000000000
B_Q01a3US			3	Grades 7-9	010000000000
B_Q01a3US			4	High school diploma	001000000000
B_Q01a3US			5	Pre-associate educat	000100000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01a3US			7	A certificate from a	000010000000
B_Q01a3US			8	Associate degree	000001000000
B_Q01a3US			9	Bachelor's degree (e	000000100000
B_Q01a3US			10	Master's degree (e.g	000000010000
B_Q01a3US			11	Professional degree	000000001000
B_Q01a3US			12	Doctorate degree (e.	000000000100
B_Q01a3US			96	Valid skip	000000000010
B_Q01aAT	19	Education - Highest qualification - Level - NATION	-1	Missing	000000000000000001
B_Q01aAT			1	No compulsory school	000000000000000000
B_Q01aAT			2	Compulsory school	100000000000000000
B_Q01aAT			3	Apprenticeship	010000000000000000
B_Q01aAT			4	Vocational School (<	001000000000000000
B_Q01aAT			5	Vocational School (2	000100000000000000
B_Q01aAT			6	Nursing	000010000000000000
B_Q01aAT			7	Master craftsman's c	000001000000000000
B_Q01aAT			8	Academic Secondary S	000000100000000000
B_Q01aAT			9	Vocational college	000000010000000000
B_Q01aAT			10	Post-secondary cours	000000001000000000
B_Q01aAT			11	Post-secondary colle	000000000100000000
B_Q01aAT			12	University courses	000000000010000000
B_Q01aAT			13	University-Bachelor	000000000001000000
B_Q01aAT			14	University-Master	000000000000100000
B_Q01aAT			15	Post-graduate course	000000000000010000
B_Q01aAT			16	Doctoral Programme	000000000000001000
B_Q01aAT			17	Foreign qualificatio	000000000000000100
B_Q01aAT			96	Valid skip	000000000000000010
B_Q01aAU1	9	Education - Highest primary/secondary school - Com	-1	Missing	00000001
B_Q01aAU1			1	Year 12 or equivalen	00000000
B_Q01aAU1			2	Year 11 or equivalen	10000000
B_Q01aAU1			3	Year 10 or equivalen	01000000
B_Q01aAU1			4	Year 9 or equivalent	00100000
B_Q01aAU1			5	Year 8 or equivalent	00010000
B_Q01aAU1			6	Year 7 or below	00001000
B_Q01aAU1			7	Never attended schoo	00000100
B_Q01aAU1			96	Valid skip	00000010
B_Q01aAU10	6	Education - Highest qualification - Completed Leve	-1	Missing	00001
B_Q01aAU10			1	Level (to be specifi	00000
B_Q01aAU10			2	Year 12 certificate	10000
B_Q01aAU10			3	Statement of attainm	01000
B_Q01aAU10			4	Foreign Qualificiati	00100
B_Q01aAU10			6	Valid skip	00010
B_Q01aAU11	4	Education - Completed any other qualifications	-1	Missing	001
B_Q01aAU11			1	Yes	000
B_Q01aAU11			2	No	100
B_Q01aAU11			6	Valid skip	010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aAU3	14	Education - Highest qualification - Month of finis	-1	Missing	0000000000001
B_Q01aAU3			1	January	0000000000000
B_Q01aAU3			2	February	1000000000000
B_Q01aAU3			3	March	0100000000000
B_Q01aAU3			4	April	0010000000000
B_Q01aAU3			5	May	0001000000000
B_Q01aAU3			6	June	0000100000000
B_Q01aAU3			7	July	0000010000000
B_Q01aAU3			8	August	0000001000000
B_Q01aAU3			9	September	0000000100000
B_Q01aAU3			10	October	0000000010000
B_Q01aAU3			11	November	0000000001000
B_Q01aAU3			12	December	0000000000100
B_Q01aAU3			96	Valid skip	0000000000010
B_Q01aAU4	4	Education - Did you complete primary school	-1	Missing	001
B_Q01aAU4			1	Yes	000
B_Q01aAU4			2	No	100
B_Q01aAU4			6	Valid skip	010
B_Q01aAU5	6	Education - Highest primary/secondary school - Cur	-1	Missing	00001
B_Q01aAU5			1	Year 12 or equivalen	00000
B_Q01aAU5			2	Year 11 or equivalen	10000
B_Q01aAU5			3	Year 10 or equivalen	01000
B_Q01aAU5			4	Year 9 or below	00100
B_Q01aAU5			6	Valid skip	00010
B_Q01aAU6	4	Education - Undertaking VET subjects/courses as pa	-1	Missing	001
B_Q01aAU6			1	Yes	000
B_Q01aAU6			2	No	100
B_Q01aAU6			6	Valid skip	010
B_Q01aAU9	4	Education - Completed trade certificate, diploma,	-1	Missing	001
B_Q01aAU9			1	Yes	000
B_Q01aAU9			2	No	100
B_Q01aAU9			6	Valid skip	010
B_Q01aBE	14	Education - Highest qualification - Level	-1	Missing	0000000000001
B_Q01aBE			1	No formal qualificat	0000000000000
B_Q01aBE			2	ISCED 1	1000000000000
B_Q01aBE			3	ISCED 2	0100000000000
B_Q01aBE			4	ISCED 3C 2 years or	0010000000000
B_Q01aBE			5	ISCED 3A-B	0001000000000
B_Q01aBE			6	ISCED 3 (without dis	0000100000000
B_Q01aBE			7	ISCED 4A-B	0000010000000
B_Q01aBE			8	ISCED 5B	0000001000000
B_Q01aBE			9	ISCED 5A, bachelor d	0000000100000
B_Q01aBE			10	ISCED 5A, master deg	0000000010000
B_Q01aBE			11	ISCED 6	0000000001000
B_Q01aBE			12	Foreign qualificatio	0000000000100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aBE			96	Valid skip	0000000000010
B_Q01aca1	4	Education - Overall education - Graduated from hig	-1	Missing	001
B_Q01aca1			1	Yes	000
B_Q01aca1			2	No	100
B_Q01aca1			6	Valid skip	010
B_Q01aca10	12	Education - Overall education - Country attained l	-1	Missing	00000000001
B_Q01aca10			1	China (People's Repu	00000000000
B_Q01aca10			2	Germany	10000000000
B_Q01aca10			3	Hong Kong	01000000000
B_Q01aca10			4	India	00100000000
B_Q01aca10			5	Italy	00010000000
B_Q01aca10			6	Jamaica	00001000000
B_Q01aca10			7	Philippines	00000100000
B_Q01aca10			8	United Kingdom (e.g.	00000010000
B_Q01aca10			9	United States	00000001000
B_Q01aca10			10	Other - specify	00000000100
B_Q01aca10			96	Valid skip	00000000010
B_Q01aca2	7	Education - Overall education - Highest grade of e	-1	Missing	000001
B_Q01aca2			1	Less than Grade 6	000000
B_Q01aca2			2	Grade 6	100000
B_Q01aca2			3	Grade 7-8 (Secondary	010000
B_Q01aca2			4	Grade 9 (Secondary 3	001000
B_Q01aca2			5	Grade 10 - 13 (Secon	000100
B_Q01aca2			6	Valid skip	000010
B_Q01aca3	17	Education - Overall education - Province/territory	-1	Missing	0000000000000001
B_Q01aca3			10	Newfoundland	0000000000000000
B_Q01aca3			11	Prince Edward Island	1000000000000000
B_Q01aca3			12	Nova Scotia	0100000000000000
B_Q01aca3			13	New Brunswick	0010000000000000
B_Q01aca3			24	Quebec	0001000000000000
B_Q01aca3			35	Ontario	0000100000000000
B_Q01aca3			46	Manitoba	0000010000000000
B_Q01aca3			47	Saskatchewan	0000001000000000
B_Q01aca3			48	Alberta	0000000100000000
B_Q01aca3			59	British Columbia	0000000010000000
B_Q01aca3			60	Yukon	0000000001000000
B_Q01aca3			61	Northwest Territorie	0000000000100000
B_Q01aca3			62	Nunavut	0000000000010000
B_Q01aca3			76	U.S.A.	0000000000001000
B_Q01aca3			77	Outside Canada/U.S.A	0000000000000100
B_Q01aca3			96	Valid skip	0000000000000010
B_Q01aca5	4	Education - Overall education - High/secondary sch	-1	Missing	001
B_Q01aca5			1	Yes	000
B_Q01aca5			2	No	100
B_Q01aca5			6	Valid skip	010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aca6	17	Education - Overall education - Highest level of s	-1	Missing	0000000000000001
B_Q01aca6			1	No formal education	0000000000000000
B_Q01aca6			2	Less than high schoo	1000000000000000
B_Q01aca6			3	High school diploma	0100000000000000
B_Q01aca6			4	Trade/vocational cer	0010000000000000
B_Q01aca6			5	Apprenticeship certi	0001000000000000
B_Q01aca6			6	CEGEP diploma or cer	0000100000000000
B_Q01aca6			7	Non-university certi	0000010000000000
B_Q01aca6			8	University transfer	0000001000000000
B_Q01aca6			9	University certifica	0000000100000000
B_Q01aca6			10	Bachelor's degree	0000000010000000
B_Q01aca6			11	University certifica	0000000001000000
B_Q01aca6			12	First professional d	0000000000100000
B_Q01aca6			13	Master's	0000000000010000
B_Q01aca6			14	Ph.D.	0000000000001000
B_Q01aca6			15	Education not defina	0000000000000100
B_Q01aca6			96	Valid skip	0000000000000010
B_Q01aca7	4	Education - Overall education - CEGEP diploma/cert	-1	Missing	001
B_Q01aca7			1	Yes	000
B_Q01aca7			2	No	100
B_Q01aca7			6	Valid skip	010
B_Q01aca8	9	Education - Overall education - Length - Complete	-1	Missing	00000001
B_Q01aca8			1	Less than 3 months	00000000
B_Q01aca8			2	3 months to less tha	10000000
B_Q01aca8			3	One year	01000000
B_Q01aca8			4	Greater than one yea	00100000
B_Q01aca8			5	Two years	00010000
B_Q01aca8			6	Greater than two yea	00001000
B_Q01aca8			7	Three years or more	00000100
B_Q01aca8			96	Valid skip	00000010
B_Q01aca9	4	Education - Overall education - Obtained trade/voc	-1	Missing	001
B_Q01aca9			1	Yes	000
B_Q01aca9			2	No	100
B_Q01aca9			6	Valid skip	010
B_Q01aCY	11	Education - Highest qualification - Level	-1	Missing	0000000001
B_Q01aCY			1	I never went to scho	0000000000
B_Q01aCY			2	Primary school	1000000000
B_Q01aCY			3	Public/Private Secon	0100000000
B_Q01aCY			4	High School/Vocation	0010000000
B_Q01aCY			5	Non-Univ. Degree/Dip	0001000000
B_Q01aCY			6	Undergraduate degree	0000100000
B_Q01aCY			7	Postgraduate degree,	0000010000
B_Q01aCY			8	Doctorate	0000001000
B_Q01aCY			9	Foreign qualificatio	0000000100
B_Q01aCY			96	Valid skip	0000000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aCZ	16	Education - Highest qualification - Level	-1	Missing	000000000000001
B_Q01aCZ			1	No formal education	000000000000000
B_Q01aCZ			2	First level of basic	100000000000000
B_Q01aCZ			3	basic ISCED 2	010000000000000
B_Q01aCZ			4	vocational without m	001000000000000
B_Q01aCZ			5	vocational without m	000100000000000
B_Q01aCZ			6	ISCED 3A vocational	000010000000000
B_Q01aCZ			7	ISCED 3A technical w	000001000000000
B_Q01aCZ			8	ISCED 3A general wit	000000100000000
B_Q01aCZ			9	ISCED 4 follow-up co	000000010000000
B_Q01aCZ			10	ISCED 5B higher prof	000000001000000
B_Q01aCZ			11	ISCED 5A, bachelor	000000000100000
B_Q01aCZ			12	ISCED 5A, master	000000000010000
B_Q01aCZ			13	ISCED 6, post gradua	000000000001000
B_Q01aCZ			14	Foreign qualificatio	000000000000100
B_Q01aCZ			96	Valid skip	000000000000010
B_Q01aDE1	14	Education National - Highest school qualification	-1	Missing	0000000000001
B_Q01aDE1			1	No formal education	0000000000000
B_Q01aDE1			2	No Hauptschulabschlu	1000000000000
B_Q01aDE1			3	Hauptschulabschluss	0100000000000
B_Q01aDE1			4	Realschulabschluss (	0010000000000
B_Q01aDE1			5	Polytechnische Obers	0001000000000
B_Q01aDE1			6	Polytechnische Obers	0000100000000
B_Q01aDE1			7	Fachhochschulreife,	0000010000000
B_Q01aDE1			8	Abitur/EOS (General	0000001000000
B_Q01aDE1			9	Abitur (General high	0000000100000
B_Q01aDE1			10	Foreign school leavi	0000000010000
B_Q01aDE1			11	Another school leavi	0000000001000
B_Q01aDE1			12	No school qualificat	0000000000100
B_Q01aDE1			96	Valid skip	0000000000010
B_Q01aDE1_REC	14	Education National - Highest school qualification	-1	Missing	0000000000001
B_Q01aDE1_REC			1	No formal education	0000000000000
B_Q01aDE1_REC			2	No Hauptschulabschlu	1000000000000
B_Q01aDE1_REC			3	Hauptschulabschluss	0100000000000
B_Q01aDE1_REC			4	Realschulabschluss (	0010000000000
B_Q01aDE1_REC			5	Polytechnische Obers	0001000000000
B_Q01aDE1_REC			6	Polytechnische Obers	0000100000000
B_Q01aDE1_REC			7	Fachhochschulreife,	0000010000000
B_Q01aDE1_REC			8	Abitur/EOS (General	0000001000000
B_Q01aDE1_REC			9	Abitur (General high	0000000100000
B_Q01aDE1_REC			10	Foreign school leavi	0000000010000
B_Q01aDE1_REC			11	Another school leavi	0000000001000
B_Q01aDE1_REC			12	No school qualificat	0000000000100
B_Q01aDE1_REC			96	Valid skip	0000000000010
B_Q01aDE2	14	Education National - Highest professional qualific	-1	Missing	0000000000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aDE2			1	No professional qual	0000000000000
B_Q01aDE2			2	Apprenticeship (Lehr	1000000000000
B_Q01aDE2			3	Basic vocational tra	0100000000000
B_Q01aDE2			4	Training at Fachschu	0010000000000
B_Q01aDE2			5	Berufsakademie, Fach	0001000000000
B_Q01aDE2			6	Bachelor at Fachhoch	0000100000000
B_Q01aDE2			7	Master/Diplom at Fac	0000010000000
B_Q01aDE2			8	Bachelor at universi	0000001000000
B_Q01aDE2			9	Master/Diplom at uni	0000000100000
B_Q01aDE2			10	Doctorate	0000000010000
B_Q01aDE2			11	Foreign professional	0000000001000
B_Q01aDE2			12	Another professional	0000000000100
B_Q01aDE2			96	Valid skip	0000000000010
B_Q01aDE2_REC	14	Education National - Highest professional qualific	-1	Missing	0000000000001
B_Q01aDE2_REC			1	No professional qual	0000000000000
B_Q01aDE2_REC			2	Apprenticeship (Lehr	1000000000000
B_Q01aDE2_REC			3	Basic vocational tra	0100000000000
B_Q01aDE2_REC			4	Training at Fachschu	0010000000000
B_Q01aDE2_REC			5	Berufsakademie, Fach	0001000000000
B_Q01aDE2_REC			6	Bachelor at Fachhoch	0000100000000
B_Q01aDE2_REC			7	Master/Diplom at Fac	0000010000000
B_Q01aDE2_REC			8	Bachelor at universi	0000001000000
B_Q01aDE2_REC			9	Master/Diplom at uni	0000000100000
B_Q01aDE2_REC			10	Doctorate	0000000010000
B_Q01aDE2_REC			11	Foreign professional	0000000001000
B_Q01aDE2_REC			12	Another professional	0000000000100
B_Q01aDE2_REC			96	Valid skip	0000000000010
B_Q01aDK	17	Education - Highest qualification - Level	-1	Missing	0000000000000001
B_Q01aDK			1	No formal education	0000000000000000
B_Q01aDK			2	Primary school, grad	1000000000000000
B_Q01aDK			3	Lower secondary, gra	0100000000000000
B_Q01aDK			4	Upper secondary voca	0010000000000000
B_Q01aDK			5	Upper secondary voca	0001000000000000
B_Q01aDK			6	Upper secondary gene	0000100000000000
B_Q01aDK			7	Upper secondary unde	0000010000000000
B_Q01aDK			8	Post secondary short	0000001000000000
B_Q01aDK			9	Post secondary entra	0000000100000000
B_Q01aDK			10	Post secondary non t	0000000010000000
B_Q01aDK			11	Tertiary not researc	0000000001000000
B_Q01aDK			12	Bachelor degree	0000000000100000
B_Q01aDK			13	Master degree	0000000000010000
B_Q01aDK			14	Ph.d or otther resea	0000000000001000
B_Q01aDK			15	Foreign qualificatio	0000000000000100
B_Q01aDK			96	Valid skip	0000000000000010
B_Q01aEE	21	Education - Highest qualification - Level	-1	Missing	00000000000000000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aEE			1	Without primary educ	0000000000000000000
B_Q01aEE			2	Primary education	1000000000000000000
B_Q01aEE			3	Basic education	0100000000000000000
B_Q01aEE			4	General secondary ed	0010000000000000000
B_Q01aEE			5	Vocational education	0001000000000000000
B_Q01aEE			6	Vocational education	0000100000000000000
B_Q01aEE			7	Vocational education	0000010000000000000
B_Q01aEE			8	Vocational secondary	0000001000000000000
B_Q01aEE			9	Secondary specialise	0000000100000000000
B_Q01aEE			10	Vocational secondary	0000000010000000000
B_Q01aEE			11	Secondary specialise	0000000001000000000
B_Q01aEE			12	Applied higher educa	0000000000100000000
B_Q01aEE			13	Bachelor's degree (3	0000000000010000000
B_Q01aEE			14	Bachelor's degree (4	0000000000001000000
B_Q01aEE			15	Higher education (st	0000000000000100000
B_Q01aEE			16	Master's degree (3+2	0000000000000010000
B_Q01aEE			17	Master's degree (4+2	0000000000000001000
B_Q01aEE			18	Doctoral degree (inc	0000000000000000100
B_Q01aEE			19	Foreign qualificatio	0000000000000000010
B_Q01aEE			96	Valid skip	0000000000000000001
B_Q01aES	14	Education - Highest qualification - Level	-1	Missing	0000000000001
B_Q01aES			1	Not stated	0000000000000
B_Q01aES			2	Not stated	1000000000000
B_Q01aES			3	Not stated	0100000000000
B_Q01aES			4	Not stated	0010000000000
B_Q01aES			5	Not stated	0001000000000
B_Q01aES			6	Bachillerato, antigu	0000100000000
B_Q01aES			7	Pruebas de acceso a	0000010000000
B_Q01aES			8	Pruebas de acceso a	0000001000000
B_Q01aES			9	Pruebas de acceso a	0000000100000
B_Q01aES			10	Pruebas de aster y e	0000000010000
B_Q01aES			11	Programas de doctora	0000000001000
B_Q01aES			12	ProgramasN EXTRANJER	0000000000100
B_Q01aES			96	Valid skip	0000000000010
B_Q01aFI	14	Education - Highest qualification - Level	-1	Missing	0000000000001
B_Q01aFI			1	No formal qualificat	0000000000000
B_Q01aFI			2	ISCED 1	1000000000000
B_Q01aFI			3	ISCED 2	0100000000000
B_Q01aFI			4	Upper secondary voca	0010000000000
B_Q01aFI			5	General upper second	0001000000000
B_Q01aFI			6	Specialist vocationa	0000100000000
B_Q01aFI			7	Vocational post-seco	0000010000000
B_Q01aFI			8	Polytechnic degree (	0000001000000
B_Q01aFI			9	Bachelor's degree (I	0000000100000
B_Q01aFI			10	Master's degree (ISC	0000000010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aFI			11	Licentiate's and doc	0000000001000
B_Q01aFI			12	Foreign qualificatio	0000000000100
B_Q01aFI			96	Valid skip	0000000000010
B_Q01aFR1	7	Education - Highest qualification - Level	-1	Missing	000001
B_Q01aFR1			1	No formal education	000000
B_Q01aFR1			2	ISCED 1	100000
B_Q01aFR1			3	ISCED 234C	010000
B_Q01aFR1			4	ISCED 4C56	001000
B_Q01aFR1			5	Foreign qualificatio	000100
B_Q01aFR1			6	Valid skip	000010
B_Q01aFR2	11	Education - Highest qualification - Level	-1	Missing	0000000001
B_Q01aFR2			1	Secondary education	0000000000
B_Q01aFR2			2	Secondary education	1000000000
B_Q01aFR2			3	Secondary education	0100000000
B_Q01aFR2			4	Secondary education	0010000000
B_Q01aFR2			5	Secondary education	0001000000
B_Q01aFR2			6	Secondary education	0000100000
B_Q01aFR2			7	Secondary education	0000010000
B_Q01aFR2			8	Secondary education	0000001000
B_Q01aFR2			9	Secondary education	0000000100
B_Q01aFR2			96	Valid skip	0000000010
B_Q01aFR3	10	Education - Highest qualification - Diploma	-1	Missing	000000001
B_Q01aFR3			1	No diploma	000000000
B_Q01aFR3			2	Primary school certi	100000000
B_Q01aFR3			3	Secondary education	010000000
B_Q01aFR3			4	Vocational training	001000000
B_Q01aFR3			5	Technological baccal	000100000
B_Q01aFR3			6	Professional baccala	000010000
B_Q01aFR3			7	Professional or tech	000001000
B_Q01aFR3			8	General baccalaur ├⌐a	000000100
B_Q01aFR3			96	Valid skip	000000010
B_Q01aFR4	10	Education - Highest qualification - Diploma	-1	Missing	000000001
B_Q01aFR4			1	ISCED 4C	000000000
B_Q01aFR4			2	ISCED 4A-B	100000000
B_Q01aFR4			3	ISCED 4 (without dis	010000000
B_Q01aFR4			4	ISCED 5B	001000000
B_Q01aFR4			5	ISCED 5A, bachelor d	000100000
B_Q01aFR4			6	ISCED 5A, master deg	000010000
B_Q01aFR4			7	ISCED 6	000001000
B_Q01aFR4			8	Other. Please specif	000000100
B_Q01aFR4			96	Valid skip	000000010
B_Q01aIE	16	Education - Highest qualification - Level	-1	Missing	000000000000001
B_Q01aIE			1	No formal education	000000000000000
B_Q01aIE			2	Primary education (o	100000000000000
B_Q01aIE			3	Secondary 1 (Junior/	010000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aIE			4	Transition year prog	001000000000000
B_Q01aIE			5	Secondary 2 (Leaving	000100000000000
B_Q01aIE			6	Technical or Vocatio	000010000000000
B_Q01aIE			7	Advanced Certificate	000001000000000
B_Q01aIE			8	Higher Certificate (	000000100000000
B_Q01aIE			9	Diploma (e.g. Nation	000000010000000
B_Q01aIE			10	Honours Bachelor Deg	000000001000000
B_Q01aIE			11	Professional (Honour	000000000100000
B_Q01aIE			12	Post-Graduate (e.g.	000000000010000
B_Q01aIE			13	Doctorate or higher	000000000001000
B_Q01aIE			14	Foreign qualificatio	000000000000100
B_Q01aIE			96	Valid skip	000000000000010
B_Q01aIT	14	Education - Highest qualification - Level	-1	Missing	0000000000001
B_Q01aIT			1	Non formal qualifica	0000000000000
B_Q01aIT			2	Primary education or	1000000000000
B_Q01aIT			3	Lower secondary or s	0100000000000
B_Q01aIT			4	Regional Vocational	0010000000000
B_Q01aIT			5	Educational and voca	0001000000000
B_Q01aIT			6	Upper secondary educ	0000100000000
B_Q01aIT			7	Post-second. non ter	0000010000000
B_Q01aIT			8	Music Conservatory D	0000001000000
B_Q01aIT			9	First stage of terti	0000000100000
B_Q01aIT			10	First or second leve	0000000010000
B_Q01aIT			11	Research Doctoral de	0000000001000
B_Q01aIT			12	Foreign qualificatio	0000000000100
B_Q01aIT			96	Valid skip	0000000000010
B_Q01aJP	17	Education - Highest qualification - Level	-1	Missing	0000000000000001
B_Q01aJP			1	No formal school edu	0000000000000000
B_Q01aJP			2	Elementary school	1000000000000000
B_Q01aJP			3	Lower secondary scho	0100000000000000
B_Q01aJP			4	Short-term course of	0010000000000000
B_Q01aJP			5	Specialized course o	0001000000000000
B_Q01aJP			6	General/integrated c	0000100000000000
B_Q01aJP			7	Passed upper seconda	0000010000000000
B_Q01aJP			8	Advanced course of u	0000001000000000
B_Q01aJP			9	Regular/Advanced cou	0000000100000000
B_Q01aJP			10	Undergraduate progra	0000000010000000
B_Q01aJP			11	Master's program/Doc	0000000001000000
B_Q01aJP			12	Completed all work o	0000000000100000
B_Q01aJP			13	Doctoral programs of	0000000000010000
B_Q01aJP			14	Specialized training	0000000000001000
B_Q01aJP			15	Foreign qualificatio	0000000000000100
B_Q01aJP			96	Valid skip	0000000000000010
B_Q01aJPX1	15	Education - Highest qualification - Level	-1	Missing	00000000000001
B_Q01aJPX1			1	No formal school edu	00000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aJPX1			2	Elementary school	10000000000000
B_Q01aJPX1			3	Lower secondary scho	01000000000000
B_Q01aJPX1			4	Short-term course of	00100000000000
B_Q01aJPX1			5	Specialized course o	00010000000000
B_Q01aJPX1			6	General/integrated c	00001000000000
B_Q01aJPX1			7	Passed upper seconda	00000100000000
B_Q01aJPX1			8	Advanced course of u	00000010000000
B_Q01aJPX1			9	Regular/advanced cou	00000001000000
B_Q01aJPX1			10	Undergraduate progra	00000000100000
B_Q01aJPX1			11	Master's program/Doc	00000000010000
B_Q01aJPX1			12	Completed all work o	00000000001000
B_Q01aJPX1			13	Doctoral programs of	00000000000100
B_Q01aJPX1			96	Valid skip	00000000000010
B_Q01aJPX2	4	Education - Highets qualification - Scholarship	-1	Missing	001
B_Q01aJPX2			1	Yes	000
B_Q01aJPX2			2	No	100
B_Q01aJPX2			6	Valid skip	010
B_Q01aKO	14	KO_Education - Highest qualification - Level	-1	Missing	0000000000001
B_Q01aKO			1	no formal education	0000000000000
B_Q01aKO			2	Elementary school	1000000000000
B_Q01aKO			3	Middle school	0100000000000
B_Q01aKO			4	High school(college	0010000000000
B_Q01aKO			5	High school(vocation	0001000000000
B_Q01aKO			6	2-3 year college	0000100000000
B_Q01aKO			7	4 year college(speci	0000010000000
B_Q01aKO			8	4 year college(gener	0000001000000
B_Q01aKO			9	Master's degree(spec	0000000100000
B_Q01aKO			10	Master's degree(gene	0000000010000
B_Q01aKO			11	Doctoral degree	0000000001000
B_Q01aKO			12	Foreign qualificatio	0000000000100
B_Q01aKO			96	Valid skip	0000000000010
B_Q01aNL	19	Education - Highest qualification - Level	-1	Missing	000000000000000001
B_Q01aNL			1	no formal qualificat	000000000000000000
B_Q01aNL			2	primary education (i	100000000000000000
B_Q01aNL			3	sec education, first	010000000000000000
B_Q01aNL			4	sec education, first	001000000000000000
B_Q01aNL			5	secondary education,	000100000000000000
B_Q01aNL			6	secondary education,	000010000000000000
B_Q01aNL			7	secondary education,	000001000000000000
B_Q01aNL			8	secondary education,	000000100000000000
B_Q01aNL			9	secondary education,	000000010000000000
B_Q01aNL			10	secondary education,	000000001000000000
B_Q01aNL			11	secondary education,	000000000100000000
B_Q01aNL			12	tertiary education,	000000000010000000
B_Q01aNL			13	tertiary education,	000000000001000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aNL			14	tertiary education,	000000000000100000
B_Q01aNL			15	tertiary education,	000000000000010000
B_Q01aNL			16	tertiary education,	000000000000001000
B_Q01aNL			17	foreign qualificatio	000000000000000100
B_Q01aNL			96	Valid skip	000000000000000010
B_Q01aNO	15	Education - Highest qualification - Level	-1	Missing	00000000000001
B_Q01aNO			1	No formal qualificat	00000000000000
B_Q01aNO			2	ISCED 1	10000000000000
B_Q01aNO			3	ISCED 2	01000000000000
B_Q01aNO			4	ISCED 3C shorter tha	00100000000000
B_Q01aNO			5	ISCED 3C 2 years or	00010000000000
B_Q01aNO			6	ISCED 3A-B	00001000000000
B_Q01aNO			7	ISCED 4C	00000100000000
B_Q01aNO			8	ISCED 4A-B	00000010000000
B_Q01aNO			9	ISCED 5B	00000001000000
B_Q01aNO			10	ISCED 5A, bachelor d	00000000100000
B_Q01aNO			11	ISCED 5A, Master deg	00000000010000
B_Q01aNO			12	ISCED 6	00000000001000
B_Q01aNO			13	Foreign qualificatio	00000000000100
B_Q01aNO			96	Valid skip	00000000000010
B_Q01aPL	13	Education - Highest qualification - Level	-1	Missing	000000000001
B_Q01aPL			1	No formal qualificat	000000000000
B_Q01aPL			2	ISCED 1	100000000000
B_Q01aPL			3	ISCED 2	010000000000
B_Q01aPL			4	ISCED 3C	001000000000
B_Q01aPL			5	ISCED 3B	000100000000
B_Q01aPL			6	ISCED 3A	000010000000
B_Q01aPL			7	ISCED 4	000001000000
B_Q01aPL			8	BA, ISCED 5A (I degr	000000100000
B_Q01aPL			9	MA, ISCED 5A (II deg	000000010000
B_Q01aPL			10	ISCED 6	000000001000
B_Q01aPL			11	Foreign qualificatio	000000000100
B_Q01aPL			96	Valid skip	000000000010
B_Q01aRU	12	Education - Highest qualification - Level	-1	Missing	00000000001
B_Q01aRU			1	No formal qualificat	00000000000
B_Q01aRU			2	ISCED 1	10000000000
B_Q01aRU			3	ISCED 2	01000000000
B_Q01aRU			4	ISCED 3 (without dis	00100000000
B_Q01aRU			5	ISCED 4 (without dis	00010000000
B_Q01aRU			6	ISCED 5B	00001000000
B_Q01aRU			7	ISCED 5A, bachelor d	00000100000
B_Q01aRU			8	ISCED 5A, master deg	00000010000
B_Q01aRU			9	ISCED 6	00000001000
B_Q01aRU			10	Foreign qualificatio	00000000100
B_Q01aRU			96	Valid skip	00000000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aSE1	19	Education - Highest qualification - Level	-1	Missing	000000000000000001
B_Q01aSE1			1	Not stated or inferr	000000000000000000
B_Q01aSE1			2	Not stated or inr	100000000000000000
B_Q01aSE1			3	Grundskola, enhetssk	010000000000000000
B_Q01aSE1			4	Yrkesutbildning	001000000000000000
B_Q01aSE1			5	Grundskolekompetens	000100000000000000
B_Q01aSE1			6	Flickskola	000010000000000000
B_Q01aSE1			7	Gymnasie fackskola y	000001000000000000
B_Q01aSE1			8	Gymnasie fackskola y	000000100000000000
B_Q01aSE1			9	Gymnasie fackskola y	000000010000000000
B_Q01aSE1			10	Vuxenutbildning mots	000000001000000000
B_Q01aSE1			11	Vuxenutbildning mots	000000000100000000
B_Q01aSE1			12	Eftergymnasial utbil	000000000010000000
B_Q01aSE1			13	Eftergymnasial utbil	000000000001000000
B_Q01aSE1			14	Eftergymnasial utbil	000000000000100000
B_Q01aSE1			15	Eftergymnasial utbil	000000000000010000
B_Q01aSE1			16	Forskarutbildning	000000000000001000
B_Q01aSE1			17	Forsndsk utbildning	000000000000000100
B_Q01aSE1			96	Valid skip	000000000000000010
B_Q01aSE2	4	Degree	-1	Missing	001
B_Q01aSE2			1	Yes	000
B_Q01aSE2			2	No	100
B_Q01aSE2			6	Valid skip	010
B_Q01aSE3	6	Type of degree	-1	Missing	00001
B_Q01aSE3			1	Fil Kand	00000
B_Q01aSE3			2	Fil Mag	10000
B_Q01aSE3			3	Master	01000
B_Q01aSE3			4	Annan typ av examen	00100
B_Q01aSE3			6	Valid skip	00010
B_Q01aSK	14	Education - Highest qualification - Level	-1	Missing	0000000000001
B_Q01aSK			1	Pre school education	0000000000000
B_Q01aSK			2	Primary school 1-4.	1000000000000
B_Q01aSK			3	Primary school 5.-9.	0100000000000
B_Q01aSK			4	Secondary technical	0010000000000
B_Q01aSK			5	Secondary technical	0001000000000
B_Q01aSK			6	Secondary schools wi	0000100000000
B_Q01aSK			7	Upper secondary scho	0000010000000
B_Q01aSK			8	Pre-tertiary school,	0000001000000
B_Q01aSK			9	Bachelor degree, Gra	0000000100000
B_Q01aSK			10	Master degree	0000000010000
B_Q01aSK			11	PhD studies, Second	0000000001000
B_Q01aSK			12	Foreign qualificatio	0000000000100
B_Q01aSK			96	Valid skip	0000000000010
B_Q01aUK1	30	Education - Highest qualification - Level	-1	Missing	00000000000000000000000000001
B_Q01aUK1			1	Degree level qualifi	00000000000000000000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aUK1			2	Diploma in higher ed	10000000000000000000000000000
B_Q01aUK1			3	HNC/HND	01000000000000000000000000000
B_Q01aUK1			4	ONC/OND	00100000000000000000000000000
B_Q01aUK1			5	BTEC, BEC, TEC or Ed	00010000000000000000000000000
B_Q01aUK1			6	SCOTVEC, SCOTEC, SCO	00001000000000000000000000000
B_Q01aUK1			7	Teaching qualificati	00000100000000000000000000000
B_Q01aUK1			8	Nursing or other med	00000010000000000000000000000
B_Q01aUK1			9	other Higher Educati	00000001000000000000000000000
B_Q01aUK1			10	A Level/Vocational A	00000000100000000000000000000
B_Q01aUK1			11	Highers (Scotland)	00000000010000000000000000000
B_Q01aUK1			12	NVQ/SVQ	00000000001000000000000000000
B_Q01aUK1			13	GNVQ/GSVQ	00000000000100000000000000000
B_Q01aUK1			14	AS Level/Vocational	00000000000010000000000000000
B_Q01aUK1			15	Advanced highers or	00000000000001000000000000000
B_Q01aUK1			16	Access to HE	00000000000000100000000000000
B_Q01aUK1			17	O Level/GCSE/Vocatio	00000000000000010000000000000
B_Q01aUK1			18	Intermediate 1 or 2	00000000000000001000000000000
B_Q01aUK1			19	Standard Grade or O	00000000000000000100000000000
B_Q01aUK1			20	National Qualificati	00000000000000000010000000000
B_Q01aUK1			21	RSA/OCR	00000000000000000001000000000
B_Q01aUK1			22	City and Guilds	00000000000000000000100000000
B_Q01aUK1			23	YT Certificate/YTP	00000000000000000000010000000
B_Q01aUK1			24	Key skills/Basic ski	00000000000000000000001000000
B_Q01aUK1			25	Entry level qualific	00000000000000000000000100000
B_Q01aUK1			26	Foreign qualificatio	00000000000000000000000010000
B_Q01aUK1			27	Any other profession	00000000000000000000000001000
B_Q01aUK1			28	No formal qualificat	00000000000000000000000000100
B_Q01aUK1			96	Valid skip	00000000000000000000000000010
B_Q01aUK10	7	Education - Highest SCOTVEC/SCOTEC/SCOTBEc qualifi	-1	Missing	000001
B_Q01aUK10			1	A higher Level (leve	000000
B_Q01aUK10			2	Full national certif	100000
B_Q01aUK10			3	A first diploma or g	010000
B_Q01aUK10			4	A first certificate	001000
B_Q01aUK10			5	Modules towards a Na	000100
B_Q01aUK10			6	Valid skip	000010
B_Q01aUK11	7	Education - Highest GNVQ/GSVQ qualification	-1	Missing	000001
B_Q01aUK11			1	Advanced level	000000
B_Q01aUK11			2	Full intermediate le	100000
B_Q01aUK11			3	Part 1 intermediate	010000
B_Q01aUK11			4	Full foundation leve	001000
B_Q01aUK11			5	Part 1 foundation le	000100
B_Q01aUK11			6	Valid skip	000010
B_Q01aUK12	6	Education - Highest RSA/OCR qualification	-1	Missing	00001
B_Q01aUK12			1	a higher diploma	00000
B_Q01aUK12			2	an advanced diploma	10000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aUK12			3	a diploma	01000
B_Q01aUK12			4	or some other RSA (i	00100
B_Q01aUK12			6	Valid skip	00010
B_Q01aUK13	5	Education - Highest City & Guilds qualification	-1	Missing	0001
B_Q01aUK13			1	Advanced craft/part	0000
B_Q01aUK13			2	craft/part 2	1000
B_Q01aUK13			3	foundation/part 1	0100
B_Q01aUK13			6	Valid skip	0010
B_Q01aUK2	7	Education - Highest full NVQ/SVQ - Level	-1	Missing	000001
B_Q01aUK2			1	Level 1	000000
B_Q01aUK2			2	Level 2	100000
B_Q01aUK2			3	Level 3	010000
B_Q01aUK2			4	Level 4	001000
B_Q01aUK2			5	Level 5	000100
B_Q01aUK2			6	Valid skip	000010
B_Q01aUK3	4	Education - Number of A Levels	-1	Missing	001
B_Q01aUK3			1	one A level (or equi	000
B_Q01aUK3			2	more than one	100
B_Q01aUK3			6	Valid skip	010
B_Q01aUK4	4	Education - Number of SCE Highers	-1	Missing	001
B_Q01aUK4			1	3 or more Highers	000
B_Q01aUK4			2	Fewer than 3 Highers	100
B_Q01aUK4			6	Valid skip	010
B_Q01aUK5	5	Education - Number of AS Levels	-1	Missing	0001
B_Q01aUK5			1	1 AS level	0000
B_Q01aUK5			2	2 or 3 AS levels	1000
B_Q01aUK5			3	4 or more AS levels	0100
B_Q01aUK5			6	Valid skip	0010
B_Q01aUK6_01	4	Education - O levels/GCSE levels- GCSE Grade C or	-1	Missing	001
B_Q01aUK6_01			1	Marked	000
B_Q01aUK6_01			2	Not marked	100
B_Q01aUK6_01			6	Valid skip	010
B_Q01aUK6_02	4	Education - O levels/GCSE levels- O level grade c	-1	Missing	001
B_Q01aUK6_02			1	Marked	000
B_Q01aUK6_02			2	Not marked	100
B_Q01aUK6_02			6	Valid skip	010
B_Q01aUK6_03	4	Education - O levels/GCSE levels- CSEs Grade 1	-1	Missing	001
B_Q01aUK6_03			1	Marked	000
B_Q01aUK6_03			2	Not marked	100
B_Q01aUK6_03			6	Valid skip	010
B_Q01aUK6_04	4	Education - O levels/GCSE levels- Standards Grade	-1	Missing	001
B_Q01aUK6_04			1	Marked	000
B_Q01aUK6_04			2	Not marked	100
B_Q01aUK6_04			6	Valid skip	010
B_Q01aUK6_05	4	Education - O levels/GCSE levels- intermediate 1 g	-1	Missing	001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01aUK6_05			1	Marked	000
B_Q01aUK6_05			2	Not marked	100
B_Q01aUK6_05			6	Valid skip	010
B_Q01aUK6_06	4	Education - O levels/GCSE levels- intermediate 2 g	-1	Missing	001
B_Q01aUK6_06			1	Marked	000
B_Q01aUK6_06			2	Not marked	100
B_Q01aUK6_06			6	Valid skip	010
B_Q01aUK6_07	4	Education - O levels/GCSE levels- none of these	-1	Missing	001
B_Q01aUK6_07			1	Marked	000
B_Q01aUK6_07			2	Not marked	100
B_Q01aUK6_07			6	Valid skip	010
B_Q01aUK7	4	Education - Number GCSE (or equiv) passes	-1	Missing	001
B_Q01aUK7			1	Fewer than 5	000
B_Q01aUK7			2	5 or more	100
B_Q01aUK7			6	Valid skip	010
B_Q01aUK8	6	Education - Maths/English GCSE (or equiv)	-1	Missing	00001
B_Q01aUK8			1	English	00000
B_Q01aUK8			2	Maths	10000
B_Q01aUK8			3	Both	01000
B_Q01aUK8			4	Neither	00100
B_Q01aUK8			6	Valid skip	00010
B_Q01aUK9	6	Education - Highest BTEC/BEC/TEC/EdExcel qualifica	-1	Missing	00001
B_Q01aUK9			1	A higher Level (leve	00000
B_Q01aUK9			2	National Certificate	10000
B_Q01aUK9			3	First Diploma or gen	01000
B_Q01aUK9			4	First certificate or	00100
B_Q01aUK9			6	Valid skip	00010
B_Q01aUS	14	Education - Highest qualification - Level	-1	Missing	0000000000001
B_Q01aUS			1	Pre-primary or no sc	0000000000000
B_Q01aUS			2	Grades 1-6	1000000000000
B_Q01aUS			3	Grades 7-9	0100000000000
B_Q01aUS			4	High school diploma	0010000000000
B_Q01aUS			5	Pre-associate educat	0001000000000
B_Q01aUS			7	A certificate from a	0000100000000
B_Q01aUS			8	Associate degree	0000010000000
B_Q01aUS			9	Bachelor's degree (e	0000001000000
B_Q01aUS			10	Master's degree (e.g	0000000100000
B_Q01aUS			11	Professional degree	0000000010000
B_Q01aUS			12	Doctorate degree (e.	0000000001000
B_Q01aUS			13	Foreign degree	0000000000100
B_Q01aUS			96	Valid skip	0000000000010
B_Q01bca1	11	Education - Highest level of schooling - Field of	-1	Missing	0000000001
B_Q01bca1			1	General programs	0000000000
B_Q01bca1			2	Teacher training and	1000000000
B_Q01bca1			3	Humanities, language	0100000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01bca1			4	Social sciences, bus	0010000000
B_Q01bca1			5	Science, mathematics	0001000000
B_Q01bca1			6	Engineering, manufac	0000100000
B_Q01bca1			7	Agriculture and vete	0000010000
B_Q01bca1			8	Health and welfare	0000001000
B_Q01bca1			9	Services	0000000100
B_Q01bca1			96	Valid skip	0000000010
B_Q01bCZ	13	Education - Highest qualification - Area of study	-1	Missing	000000000001
B_Q01bCZ			1	General programmes	000000000000
B_Q01bCZ			2	Teacher training and	100000000000
B_Q01bCZ			3	Humanities, language	010000000000
B_Q01bCZ			4	Social sciences	001000000000
B_Q01bCZ			5	Business and law	000100000000
B_Q01bCZ			6	Science, mathematics	000010000000
B_Q01bCZ			7	Engineering, manufac	000001000000
B_Q01bCZ			8	Agriculture and vete	000000100000
B_Q01bCZ			9	Health	000000010000
B_Q01bCZ			10	Welfare	000000001000
B_Q01bCZ			11	Services	000000000100
B_Q01bCZ			96	Valid skip	000000000010
B_Q01bKO	12	KO_Education - major	-1	Missing	00000000001
B_Q01bKO			1	General programmes	00000000000
B_Q01bKO			2	Teacher training and	10000000000
B_Q01bKO			3	Humanities, language	01000000000
B_Q01bKO			4	Social sciences, bus	00100000000
B_Q01bKO			5	Science, mathematics	00010000000
B_Q01bKO			6	Engineering, manufac	00001000000
B_Q01bKO			7	Agriculture and vete	00000100000
B_Q01bKO			8	Dental and medicine	00000010000
B_Q01bKO			9	Health and wellfare	00000001000
B_Q01bKO			10	Services	00000000100
B_Q01bKO			96	Valid skip	00000000010
B_Q01bNL	13	Education - Highest qualification - Area of study	-1	Missing	000000000001
B_Q01bNL			1	general programmes	000000000000
B_Q01bNL			2	teacher training, ed	100000000000
B_Q01bNL			3	humanities, language	010000000000
B_Q01bNL			4	social sciences, com	001000000000
B_Q01bNL			5	economy, business, m	000100000000
B_Q01bNL			6	law, civil service,	000010000000
B_Q01bNL			7	mathematics, natural	000001000000
B_Q01bNL			8	technics	000000100000
B_Q01bNL			9	agriculture, veterin	000000010000
B_Q01bNL			10	health, welfare, per	000000001000
B_Q01bNL			11	tourism, horeca, tra	000000000100
B_Q01bNL			96	Valid skip	000000000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01bUK	21	Education - Highest qualification - Area of study	-1	Missing	00000000000000000001
B_Q01bUK			1	General programmes	00000000000000000000
B_Q01bUK			2	Medicine	10000000000000000000
B_Q01bUK			3	Medical related subj	01000000000000000000
B_Q01bUK			4	Biological Sciences	00100000000000000000
B_Q01bUK			5	Agricultural science	00010000000000000000
B_Q01bUK			6	Physical/Environment	00001000000000000000
B_Q01bUK			7	Mathematical Science	00000100000000000000
B_Q01bUK			8	Engineering	00000010000000000000
B_Q01bUK			9	Technology	00000001000000000000
B_Q01bUK			10	Architecture and rel	00000000100000000000
B_Q01bUK			11	Social Sciences (inc	00000000010000000000
B_Q01bUK			12	Business and Financi	00000000001000000000
B_Q01bUK			13	Librarianship and In	00000000000100000000
B_Q01bUK			14	Linguistics, English	00000000000010000000
B_Q01bUK			15	European Languages	00000000000001000000
B_Q01bUK			16	Other languages	00000000000000100000
B_Q01bUK			17	Humanities	00000000000000010000
B_Q01bUK			18	Arts	00000000000000001000
B_Q01bUK			19	Education	00000000000000000100
B_Q01bUK			96	Valid skip	00000000000000000010
B_Q01dca2	18	Education - Highest level of education - Attained	-1	Missing	00000000000000001
B_Q01dca2			1	No Formal Education	00000000000000000
B_Q01dca2			2	Some elementary scho	10000000000000000
B_Q01dca2			3	Some high school	01000000000000000
B_Q01dca2			4	High school diploma	00100000000000000
B_Q01dca2			5	Some trade/vocationa	00010000000000000
B_Q01dca2			6	Trade/vocational cer	00001000000000000
B_Q01dca2			7	Apprenticeship certi	00000100000000000
B_Q01dca2			8	Non-university certi	00000010000000000
B_Q01dca2			9	University transfer	00000001000000000
B_Q01dca2			10	University certifica	00000000100000000
B_Q01dca2			11	Bachelor's degree	00000000010000000
B_Q01dca2			12	University certifica	00000000001000000
B_Q01dca2			13	First professional d	00000000000100000
B_Q01dca2			14	Master's	00000000000010000
B_Q01dca2			15	Ph.D.	00000000000001000
B_Q01dca2			16	Education not defina	00000000000000100
B_Q01dca2			96	Valid skip	00000000000000010
B_Q01dca3	12	Education - Highest level of education - Country	-1	Missing	00000000001
B_Q01dca3			1	China (People's Repu	00000000000
B_Q01dca3			2	Germany	10000000000
B_Q01dca3			3	Hong Kong	01000000000
B_Q01dca3			4	India	00100000000
B_Q01dca3			5	Italy	00010000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q01dca3			6	Jamaica	00001000000
B_Q01dca3			7	Philippines	00000100000
B_Q01dca3			8	United Kingdom (e.g.	00000010000
B_Q01dca3			9	United States	00000001000
B_Q01dca3			10	Other - specify	00000000100
B_Q01dca3			96	Valid skip	00000000010
B_Q01dUKX	4	Education - Highest qualification - completed an a	-1	Missing	001
B_Q01dUKX			1	Yes	000
B_Q01dUKX			2	No	100
B_Q01dUKX			6	Valid skip	010
B_Q01eJPX	10	Education - Years spent in Kindergarten	-1	Missing	000000001
B_Q01eJPX			1	Never	000000000
B_Q01eJPX			2	Less than 6 months	100000000
B_Q01eJPX			3	6 months to 1 year	010000000
B_Q01eJPX			4	1 to 1 1/2 year	001000000
B_Q01eJPX			5	1 1/2 to 2 years	000100000
B_Q01eJPX			6	2 to 2 1/2 years	000010000
B_Q01eJPX			7	2 1/2 to 3 years	000001000
B_Q01eJPX			8	3 years	000000100
B_Q01eJPX			96	Valid skip	000000010
B_Q02aAT	4	Education - Current qualification	-1	Missing	001
B_Q02aAT			1	Yes	000
B_Q02aAT			2	No	100
B_Q02aAT			6	Valid skip	010
B_Q02aAU	4	Education - Currently studying	-1	Missing	001
B_Q02aAU			1	Yes	000
B_Q02aAU			2	No	100
B_Q02aAU			6	Valid skip	010
B_Q02aDE	6	Education National - Current qualification	-1	Missing	00001
B_Q02aDE			1	Yes, school providin	00000
B_Q02aDE			2	Yes, professional tr	10000
B_Q02aDE			3	Yes, both of the abo	01000
B_Q02aDE			4	No	00100
B_Q02aDE			6	Valid skip	00010
B_Q02aIEX	10	Education - Reason for early school leaving	-1	Missing	000000001
B_Q02aIEX			1	Had enough education	000000000
B_Q02aIEX			2	Had to work\financia	100000000
B_Q02aIEX			3	Wanted to work \ wan	010000000
B_Q02aIEX			4	Family reasons (e.g.	001000000
B_Q02aIEX			5	Did not like school	000100000
B_Q02aIEX			6	Did not do well in s	000010000
B_Q02aIEX			7	Personal illness or	000001000
B_Q02aIEX			8	School not available	000000100
B_Q02aIEX			9	Other	000000010
B_Q02b2RU	9	Education - Current qualification - Country of fo	-1	Missing	00000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02b2RU			1	Country 1	00000000
B_Q02b2RU			2	Country 2	10000000
B_Q02b2RU			3	Country 3	01000000
B_Q02b2RU			4	Country 4	00100000
B_Q02b2RU			5	Country 5	00010000
B_Q02b2RU			6	Country 6	00001000
B_Q02b2RU			7	Other country	00000100
B_Q02b2RU			96	Valid skip	00000010
B_Q02bAT	20	Education - Current qualification - Level - NATION	-1	Missing	0000000000000000001
B_Q02bAT			1	Lower secondary Scho	0000000000000000000
B_Q02bAT			2	Prevocational School	1000000000000000000
B_Q02bAT			3	Apprenticeship	0100000000000000000
B_Q02bAT			4	Vocational School (<	0010000000000000000
B_Q02bAT			5	Vocational School (2	0001000000000000000
B_Q02bAT			6	Nursing	0000100000000000000
B_Q02bAT			7	Master craftsman's c	0000010000000000000
B_Q02bAT			8	Academic secondary s	0000001000000000000
B_Q02bAT			9	1-3rd Class in a Voc	0000000100000000000
B_Q02bAT			10	4 or 5th Class in a	0000000010000000000
B_Q02bAT			11	Post-secondary cours	0000000001000000000
B_Q02bAT			12	Post-secondary colle	0000000000100000000
B_Q02bAT			13	University courses	0000000000010000000
B_Q02bAT			14	University-Bachelor	0000000000001000000
B_Q02bAT			15	University-Master	0000000000000100000
B_Q02bAT			16	Post-graduate course	0000000000000010000
B_Q02bAT			17	Doctoral Programme	0000000000000001000
B_Q02bAT			18	Foreign Qualificatio	0000000000000000100
B_Q02bAT			96	Valid skip	0000000000000000010
B_Q02bAU1	5	Education - Current qualification - Currently Stud	-1	Missing	0001
B_Q02bAU1			1	Level	0000
B_Q02bAU1			2	Year 12 or equivalen	1000
B_Q02bAU1			3	Statement of attainm	0100
B_Q02bAU1			6	Valid skip	0010
B_Q02bBE	13	Education - Current qualification - Level	-1	Missing	000000000001
B_Q02bBE			1	ISCED 1	000000000000
B_Q02bBE			2	ISCED 2	100000000000
B_Q02bBE			3	ISCED 3C 2 years or	010000000000
B_Q02bBE			4	ISCED 3A-B	001000000000
B_Q02bBE			5	ISCED 3 (without dis	000100000000
B_Q02bBE			6	ISCED 4A-B	000010000000
B_Q02bBE			7	ISCED 5B	000001000000
B_Q02bBE			8	ISCED 5A, bachelor d	000000100000
B_Q02bBE			9	ISCED 5A, master deg	000000010000
B_Q02bBE			10	ISCED 6	000000001000
B_Q02bBE			11	Foreign qualificatio	000000000100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02bBE			96	Valid skip	000000000010
B_Q02bca1	16	Education - Current study - Level of education	-1	Missing	000000000000001
B_Q02bca1			1	Grade 6	000000000000000
B_Q02bca1			2	Less than high schoo	100000000000000
B_Q02bca1			3	High school diploma	010000000000000
B_Q02bca1			4	Trade/vocational cer	001000000000000
B_Q02bca1			5	Apprenticeship certi	000100000000000
B_Q02bca1			6	CEGEP diploma or cer	000010000000000
B_Q02bca1			7	Non-university certi	000001000000000
B_Q02bca1			8	University transfer	000000100000000
B_Q02bca1			9	University certifica	000000010000000
B_Q02bca1			10	Bachelor's degree	000000001000000
B_Q02bca1			11	University certifica	000000000100000
B_Q02bca1			12	First professional d	000000000010000
B_Q02bca1			13	Master's	000000000001000
B_Q02bca1			14	Ph.D.	000000000000100
B_Q02bca1			96	Valid skip	000000000000010
B_Q02bca2	4	Education - Current study - CEGEP diploma/certific	-1	Missing	001
B_Q02bca2			1	Yes	000
B_Q02bca2			2	No	100
B_Q02bca2			6	Valid skip	010
B_Q02bca3	9	Education - Current study - Length - Complete trad	-1	Missing	00000001
B_Q02bca3			1	Less than 3 months	00000000
B_Q02bca3			2	3 months to less tha	10000000
B_Q02bca3			3	One year	01000000
B_Q02bca3			4	Greater than one yea	00100000
B_Q02bca3			5	Two years	00010000
B_Q02bca3			6	Greater than two yea	00001000
B_Q02bca3			7	Three years or more	00000100
B_Q02bca3			96	Valid skip	00000010
B_Q02bCY	9	Education - Current qualification - Level	-1	Missing	00000001
B_Q02bCY			1	Primary school	00000000
B_Q02bCY			2	Public/Private Secon	10000000
B_Q02bCY			3	High School/Vocation	01000000
B_Q02bCY			4	Non-Univ. Degree/Dip	00100000
B_Q02bCY			5	Undergraduate degree	00010000
B_Q02bCY			6	Postgraduate degree,	00001000
B_Q02bCY			7	Doctorate	00000100
B_Q02bCY			96	Valid skip	00000010
B_Q02bCZ	14	Education - Current qualification - Level	-1	Missing	0000000000001
B_Q02bCZ			1	First level of basic	0000000000000
B_Q02bCZ			2	basic ISCED 2	1000000000000
B_Q02bCZ			3	vocational without m	0100000000000
B_Q02bCZ			4	vocational without m	0010000000000
B_Q02bCZ			5	ISCED 3A vocational	0001000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02bCZ			6	ISCED 3A technical w	0000100000000
B_Q02bCZ			7	ISCED 3A general wit	0000010000000
B_Q02bCZ			8	ISCED 4 follow-up co	0000001000000
B_Q02bCZ			9	ISCED 5B higher prof	0000000100000
B_Q02bCZ			10	ISCED 5A, bachelor	0000000010000
B_Q02bCZ			11	ISCED 5A, master	0000000001000
B_Q02bCZ			12	ISCED 6, post gradua	0000000000100
B_Q02bCZ			96	Valid skip	0000000000010
B_Q02bDE1	8	Education National - Current school qualification	-1	Missing	0000001
B_Q02bDE1			1	Hauptschulabschluss	0000000
B_Q02bDE1			2	Realschulabschluss (	1000000
B_Q02bDE1			3	Fachhochschulreife,	0100000
B_Q02bDE1			4	Abitur/EOS (General	0010000
B_Q02bDE1			5	Abitur (General high	0001000
B_Q02bDE1			6	Another school leavi	0000100
B_Q02bDE1			96	Valid skip	0000010
B_Q02bDE1_REC	8	Education National - Current school qualification	-1	Missing	0000001
B_Q02bDE1_REC			1	Hauptschulabschluss	0000000
B_Q02bDE1_REC			2	Realschulabschluss (	1000000
B_Q02bDE1_REC			3	Fachhochschulreife,	0100000
B_Q02bDE1_REC			4	Abitur/EOS (General	0010000
B_Q02bDE1_REC			5	Abitur (General high	0001000
B_Q02bDE1_REC			6	Another school leavi	0000100
B_Q02bDE1_REC			96	Valid skip	0000010
B_Q02bDE2	12	Education National - Current professional qualific	-1	Missing	00000000001
B_Q02bDE2			1	Completed Apprentice	00000000000
B_Q02bDE2			2	Basic vocational tra	10000000000
B_Q02bDE2			3	Training at Fachschu	01000000000
B_Q02bDE2			4	Berufsakademie, Fach	00100000000
B_Q02bDE2			5	Bachelor at Fachhoch	00010000000
B_Q02bDE2			6	Master/Diplom at Fac	00001000000
B_Q02bDE2			7	Bachelor at universi	00000100000
B_Q02bDE2			8	Master/Diplom at uni	00000010000
B_Q02bDE2			9	Doctorate	00000001000
B_Q02bDE2			10	Another professional	00000000100
B_Q02bDE2			96	Valid skip	00000000010
B_Q02bDE2_REC	12	Education National - Current professional qualific	-1	Missing	00000000001
B_Q02bDE2_REC			1	Completed Apprentice	00000000000
B_Q02bDE2_REC			2	Basic vocational tra	10000000000
B_Q02bDE2_REC			3	Training at Fachschu	01000000000
B_Q02bDE2_REC			4	Berufsakademie, Fach	00100000000
B_Q02bDE2_REC			5	Bachelor at Fachhoch	00010000000
B_Q02bDE2_REC			6	Master/Diplom at Fac	00001000000
B_Q02bDE2_REC			7	Bachelor at universi	00000100000
B_Q02bDE2_REC			8	Master/Diplom at uni	00000010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02bDE2_REC			9	Doctorate	00000001000
B_Q02bDE2_REC			10	Another professional	00000000100
B_Q02bDE2_REC			96	Valid skip	00000000010
B_Q02bDK	15	What is the level of the qualification you are cur	-1	Missing	000000000000001
B_Q02bDK			1	Primary school, grad	000000000000000
B_Q02bDK			2	Lower secondary, gra	100000000000000
B_Q02bDK			3	Upper secondary voca	010000000000000
B_Q02bDK			4	Upper secondary voca	001000000000000
B_Q02bDK			5	Upper secondary gene	000100000000000
B_Q02bDK			6	Upper secondary unde	000010000000000
B_Q02bDK			7	Post secondary short	000001000000000
B_Q02bDK			8	Post secondary entra	000000100000000
B_Q02bDK			9	Post secondary non t	000000010000000
B_Q02bDK			10	Tertiary not researc	000000001000000
B_Q02bDK			11	Bachelor degree	000000000100000
B_Q02bDK			12	Master degree	000000000010000
B_Q02bDK			13	Ph.d or otther resea	000000000001000
B_Q02bDK			96	Valid skip	000000000000100
B_Q02bEE	14	Education - Current qualification - Level	-1	Missing	00000000000001
B_Q02bEE			1	Primary education (1	00000000000000
B_Q02bEE			2	Basic education (7-9	10000000000000
B_Q02bEE			3	General secondary ed	01000000000000
B_Q02bEE			4	Vocational education	00100000000000
B_Q02bEE			5	Vocational education	00010000000000
B_Q02bEE			6	Voc ed on the basis	00001000000000
B_Q02bEE			7	Vocational secondary	00000100000000
B_Q02bEE			8	Vocational secondary	00000010000000
B_Q02bEE			9	Applied higher educa	00000001000000
B_Q02bEE			10	Bachelor's degree (3	00000000100000
B_Q02bEE			11	Master's degree (3+2	00000000010000
B_Q02bEE			12	Doctoral degree	00000000001000
B_Q02bEE			96	Valid skip	00000000000100
B_Q02bES	12	Education - Current qualification - Level	-1	Missing	000000000001
B_Q02bES			1	Not stated or inferr	000000000000
B_Q02bES			2	Not stated or inferr	100000000000
B_Q02bES			3	Not stated or inferr	010000000000
B_Q02bES			4	Not stated or inferr	001000000000
B_Q02bES			5	Bachillerato,. Y sim	000100000000
B_Q02bES			6	Pruebas de acceso a	000010000000
B_Q02bES			7	Pruebas de acceso a	000001000000
B_Q02bES			8	Pruebas de acceso a	000000100000
B_Q02bES			9	Pruebas de aster y e	000000010000
B_Q02bES			10	Programas de doctora	000000001000
B_Q02bES			96	Valid skip	000000000100
B_Q02bFI	12	Education - Current qualification - Level	-1	Missing	000000000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02bFI			1	ISCED 1	00000000000
B_Q02bFI			2	ISCED 2	10000000000
B_Q02bFI			3	Upper secondary voca	01000000000
B_Q02bFI			4	General upper second	00100000000
B_Q02bFI			5	Specialist vocationa	00010000000
B_Q02bFI			6	Vocational post-seco	00001000000
B_Q02bFI			7	Polytechnic degree (	00000100000
B_Q02bFI			8	Bachelor's degree (I	00000010000
B_Q02bFI			9	Master's degree (ISC	00000001000
B_Q02bFI			10	Licentiate's and doc	00000000100
B_Q02bFI			96	Valid skip	00000000010
B_Q02bFR1	16	Education - Current qualification - Level	-1	Missing	0000000000000001
B_Q02bFR1			1	ISCED 1	0000000000000000
B_Q02bFR1			2	ISCED 2	1000000000000000
B_Q02bFR1			3	ISCED 3C shorter tha	0100000000000000
B_Q02bFR1			4	ISCED 3C 2 years or	0010000000000000
B_Q02bFR1			5	ISCED 3A-B	0001000000000000
B_Q02bFR1			6	ISCED 3 (without dis	0000100000000000
B_Q02bFR1			7	ISCED 4C	0000010000000000
B_Q02bFR1			8	ISCED 4A-B	0000001000000000
B_Q02bFR1			9	ISCED 4 (without dis	0000000100000000
B_Q02bFR1			10	ISCED 5B	0000000010000000
B_Q02bFR1			11	ISCED 5A, bachelor d	0000000001000000
B_Q02bFR1			12	ISCED 5A, master deg	0000000000100000
B_Q02bFR1			13	ISCED 6	0000000000010000
B_Q02bFR1			14	Foreign qualificatio	0000000000001000
B_Q02bFR1			96	Valid skip	0000000000000010
B_Q02bIE	15	Education - Current qualification - Level	-1	Missing	000000000000001
B_Q02bIE			1	No formal education	000000000000000
B_Q02bIE			2	Primary education (o	100000000000000
B_Q02bIE			3	Secondary 1 (Junior/	010000000000000
B_Q02bIE			4	Transition year prog	001000000000000
B_Q02bIE			5	Secondary 2 (Leaving	000100000000000
B_Q02bIE			6	Technical or Vocatio	000010000000000
B_Q02bIE			7	Advanced Certificate	000001000000000
B_Q02bIE			8	Higher Certificate (	000000100000000
B_Q02bIE			9	Diploma (e.g. Nation	000000010000000
B_Q02bIE			10	Honours Bachelor Deg	000000001000000
B_Q02bIE			11	Professional (Honour	000000000100000
B_Q02bIE			12	Post-Graduate (e.g.	000000000010000
B_Q02bIE			13	Doctorate or higher	000000000001000
B_Q02bIE			96	Valid skip	000000000000010
B_Q02bIT	12	Education - Current qualification - Level	-1	Missing	000000000001
B_Q02bIT			1	Primary education or	000000000000
B_Q02bIT			2	Lower secondary or s	100000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02bIT			3	Regional Vocational	01000000000
B_Q02bIT			4	Educational and voca	00100000000
B_Q02bIT			5	Upper secondary educ	00010000000
B_Q02bIT			6	Post-second. non ter	00001000000
B_Q02bIT			7	Music Conservatory D	00000100000
B_Q02bIT			8	First stage of terti	00000010000
B_Q02bIT			9	First or second leve	00000001000
B_Q02bIT			10	Research Doctoral de	00000000100
B_Q02bIT			96	Valid skip	00000000010
B_Q02bJP	14	Education - Current qualification - Level	-1	Missing	0000000000001
B_Q02bJP			1	Elementary school	0000000000000
B_Q02bJP			2	Lower secondary scho	1000000000000
B_Q02bJP			3	Short-term course of	0100000000000
B_Q02bJP			4	Specialized course o	0010000000000
B_Q02bJP			5	General/integrated c	0001000000000
B_Q02bJP			6	Passed upper seconda	0000100000000
B_Q02bJP			7	Advanced course of u	0000010000000
B_Q02bJP			8	Regular/advanced cou	0000001000000
B_Q02bJP			9	Undergraduate progra	0000000100000
B_Q02bJP			10	Master's program/Doc	0000000010000
B_Q02bJP			11	Doctoral programs of	0000000001000
B_Q02bJP			12	Specialized training	0000000000100
B_Q02bJP			96	Valid skip	0000000000010
B_Q02bKO	12	KO_Education - Current education	-1	Missing	00000000001
B_Q02bKO			1	Elementary school	00000000000
B_Q02bKO			2	Middle school	10000000000
B_Q02bKO			3	High school(college	01000000000
B_Q02bKO			4	High school(vocation	00100000000
B_Q02bKO			5	2-3 year college	00010000000
B_Q02bKO			6	4 year college(speci	00001000000
B_Q02bKO			7	4 year college(gener	00000100000
B_Q02bKO			8	Master's degree(spec	00000010000
B_Q02bKO			9	Master's degree(gene	00000001000
B_Q02bKO			10	Doctoral degree	00000000100
B_Q02bKO			96	Valid skip	00000000010
B_Q02bNL	17	Education - Current qualification - Level	-1	Missing	0000000000000001
B_Q02bNL			1	primary education (i	0000000000000000
B_Q02bNL			2	sec education,first	1000000000000000
B_Q02bNL			3	sec education, first	0100000000000000
B_Q02bNL			4	secondary education,	0010000000000000
B_Q02bNL			5	secondary education,	0001000000000000
B_Q02bNL			6	secondary education,	0000100000000000
B_Q02bNL			7	secondary education,	0000010000000000
B_Q02bNL			8	secondary education,	0000001000000000
B_Q02bNL			9	sec education, secon	0000000100000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02bNL			10	secondary education,	0000000010000000
B_Q02bNL			11	tertiary education,	0000000001000000
B_Q02bNL			12	tertiary education,	0000000000100000
B_Q02bNL			13	tertiary education,	0000000000010000
B_Q02bNL			14	tertiary education,	0000000000001000
B_Q02bNL			15	tertiary education,	0000000000000100
B_Q02bNL			96	Valid skip	0000000000000010
B_Q02bNO	13	Education - Current qualification - Level	-1	Missing	000000000001
B_Q02bNO			1	ISCED 1	000000000000
B_Q02bNO			2	ISCED 2	100000000000
B_Q02bNO			3	ISCED 3C shorter tha	010000000000
B_Q02bNO			4	ISCED 3C 2 years or	001000000000
B_Q02bNO			5	ISCED 3A-B	000100000000
B_Q02bNO			6	ISCED 4C	000010000000
B_Q02bNO			7	ISCED 4A-B	000001000000
B_Q02bNO			8	ISCED 5B	000000100000
B_Q02bNO			9	ISCED 5A, bachelor d	000000010000
B_Q02bNO			10	ISCED 5A, Master deg	000000001000
B_Q02bNO			11	ISCED 6	000000000100
B_Q02bNO			96	Valid skip	000000000010
B_Q02bPL	11	Education - Current qualification - Level	-1	Missing	0000000001
B_Q02bPL			1	ISCED 1	0000000000
B_Q02bPL			2	ISCED 2	1000000000
B_Q02bPL			3	ISCED 3C	0100000000
B_Q02bPL			4	ISCED 3B	0010000000
B_Q02bPL			5	ISCED 3A	0001000000
B_Q02bPL			6	ISCED 4	0000100000
B_Q02bPL			7	BA, ISCED 5A (I degr	0000010000
B_Q02bPL			8	MA, ISCED 5A (II deg	0000001000
B_Q02bPL			9	ISCED 6	0000000100
B_Q02bPL			96	Valid skip	0000000010
B_Q02bRU	10	Education - Current qualification - Level	-1	Missing	000000001
B_Q02bRU			1	ISCED 1	000000000
B_Q02bRU			2	ISCED 2	100000000
B_Q02bRU			3	ISCED 3 (without dis	010000000
B_Q02bRU			4	ISCED 4 (without dis	001000000
B_Q02bRU			5	ISCED 5B	000100000
B_Q02bRU			6	ISCED 5A, bachelor d	000010000
B_Q02bRU			7	ISCED 5A, master deg	000001000
B_Q02bRU			8	ISCED 6	000000100
B_Q02bRU			96	Valid skip	000000010
B_Q02bSE	15	Level of education	-1	Missing	00000000000001
B_Q02bSE			1	Not stated ok 1-6	00000000000000
B_Q02bSE			2	Not stated ok 7-9	10000000000000
B_Q02bSE			3	Grundskolekompetens	01000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02bSE			4	Gymnasie fackskola y	00100000000000
B_Q02bSE			5	Gymnasie fackskola y	00010000000000
B_Q02bSE			6	Gymnasie fackskola y	00001000000000
B_Q02bSE			7	Vuxenutbildning mots	00000100000000
B_Q02bSE			8	Vuxenutbildning mots	00000010000000
B_Q02bSE			9	Eftergymnasial utbil	00000001000000
B_Q02bSE			10	Eftergymnasial utbil	00000000100000
B_Q02bSE			11	Eftergymnasial utbil	00000000010000
B_Q02bSE			12	Eftergymnasial utbil	00000000001000
B_Q02bSE			13	Forskarutbildning	00000000000100
B_Q02bSE			96	Valid skip	00000000000010
B_Q02bSK	12	Education - Current qualification - Level	-1	Missing	00000000001
B_Q02bSK			1	Primary school 1-4.	00000000000
B_Q02bSK			2	Primary school 5.-9.	10000000000
B_Q02bSK			3	Secondary technical	01000000000
B_Q02bSK			4	Secondary technical	00100000000
B_Q02bSK			5	Secondary schools wi	00010000000
B_Q02bSK			6	Upper secondary scho	00001000000
B_Q02bSK			7	Pre-tertiary school,	00000100000
B_Q02bSK			8	Bachelor degree, Gra	00000010000
B_Q02bSK			9	Master degree	00000001000
B_Q02bSK			10	PhD studies, Second	00000000100
B_Q02bSK			96	Valid skip	00000000010
B_Q02bUK1_01	4	Education - Current qualification - Level - Degree	-1	Missing	001
B_Q02bUK1_01			1	Marked	000
B_Q02bUK1_01			2	Not marked	100
B_Q02bUK1_01			6	Valid skip	010
B_Q02bUK1_02	4	Education - Current qualification - Level - Diplom	-1	Missing	001
B_Q02bUK1_02			1	Marked	000
B_Q02bUK1_02			2	Not marked	100
B_Q02bUK1_02			6	Valid skip	010
B_Q02bUK1_03	4	Education - Current qualification - Level - HNC/HN	-1	Missing	001
B_Q02bUK1_03			1	Marked	000
B_Q02bUK1_03			2	Not marked	100
B_Q02bUK1_03			6	Valid skip	010
B_Q02bUK1_04	4	Education - Current qualification - Level - ONC/ON	-1	Missing	001
B_Q02bUK1_04			1	Marked	000
B_Q02bUK1_04			2	Not marked	100
B_Q02bUK1_04			6	Valid skip	010
B_Q02bUK1_05	4	Education - Current qualification - Level - BTEC/E	-1	Missing	001
B_Q02bUK1_05			1	Marked	000
B_Q02bUK1_05			2	Not marked	100
B_Q02bUK1_05			6	Valid skip	010
B_Q02bUK1_06	4	Education - Current qualification - Level - SCOTVE	-1	Missing	001
B_Q02bUK1_06			1	Marked	000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02bUK1_06			2	Not marked	100
B_Q02bUK1_06			6	Valid skip	010
B_Q02bUK1_07	4	Education - Current qualification - Level - Teachi	-1	Missing	001
B_Q02bUK1_07			1	Marked	000
B_Q02bUK1_07			2	Not marked	100
B_Q02bUK1_07			6	Valid skip	010
B_Q02bUK1_08	4	Education - Current qualification - Level - Nursin	-1	Missing	001
B_Q02bUK1_08			1	Marked	000
B_Q02bUK1_08			2	Not marked	100
B_Q02bUK1_08			6	Valid skip	010
B_Q02bUK1_09	4	Education - Current qualification - Level - Other	-1	Missing	001
B_Q02bUK1_09			1	Marked	000
B_Q02bUK1_09			2	Not marked	100
B_Q02bUK1_09			6	Valid skip	010
B_Q02bUK1_10	4	Education - Current qualification - Level - A Leve	-1	Missing	001
B_Q02bUK1_10			1	Marked	000
B_Q02bUK1_10			2	Not marked	100
B_Q02bUK1_10			6	Valid skip	010
B_Q02bUK1_11	4	Education - Current qualification - Level - NVQ/SV	-1	Missing	001
B_Q02bUK1_11			1	Marked	000
B_Q02bUK1_11			2	Not marked	100
B_Q02bUK1_11			6	Valid skip	010
B_Q02bUK1_12	4	Education - Current qualification - Level - AS Lev	-1	Missing	001
B_Q02bUK1_12			1	Marked	000
B_Q02bUK1_12			2	Not marked	100
B_Q02bUK1_12			6	Valid skip	010
B_Q02bUK1_13	4	Education - Current qualification - Level - Access	-1	Missing	001
B_Q02bUK1_13			1	Marked	000
B_Q02bUK1_13			2	Not marked	100
B_Q02bUK1_13			6	Valid skip	010
B_Q02bUK1_14	4	Education - Current qualification - Level - Advanc	-1	Missing	001
B_Q02bUK1_14			1	Marked	000
B_Q02bUK1_14			2	Not marked	100
B_Q02bUK1_14			6	Valid skip	010
B_Q02bUK1_15	4	Education - Current qualification - Level - Higher	-1	Missing	001
B_Q02bUK1_15			1	Marked	000
B_Q02bUK1_15			2	Not marked	100
B_Q02bUK1_15			6	Valid skip	010
B_Q02bUK1_16	4	Education - Current qualification - Level - Interm	-1	Missing	001
B_Q02bUK1_16			1	Marked	000
B_Q02bUK1_16			2	Not marked	100
B_Q02bUK1_16			6	Valid skip	010
B_Q02bUK1_17	4	Education - Current qualification - Level - Interm	-1	Missing	001
B_Q02bUK1_17			1	Marked	000
B_Q02bUK1_17			2	Not marked	100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02bUK1_17			6	Valid skip	010
B_Q02bUK1_18	4	Education - Current qualification - Level - Access	-1	Missing	001
B_Q02bUK1_18			1	Marked	000
B_Q02bUK1_18			2	Not marked	100
B_Q02bUK1_18			6	Valid skip	010
B_Q02bUK1_19	4	Education - Current qualification - Level - Nation	-1	Missing	001
B_Q02bUK1_19			1	Marked	000
B_Q02bUK1_19			2	Not marked	100
B_Q02bUK1_19			6	Valid skip	010
B_Q02bUK1_20	4	Education - Current qualification - Level - GCSE/V	-1	Missing	001
B_Q02bUK1_20			1	Marked	000
B_Q02bUK1_20			2	Not marked	100
B_Q02bUK1_20			6	Valid skip	010
B_Q02bUK1_21	4	Education - Current qualification - Level - RSA/OC	-1	Missing	001
B_Q02bUK1_21			1	Marked	000
B_Q02bUK1_21			2	Not marked	100
B_Q02bUK1_21			6	Valid skip	010
B_Q02bUK1_22	4	Education - Current qualification - Level - City a	-1	Missing	001
B_Q02bUK1_22			1	Marked	000
B_Q02bUK1_22			2	Not marked	100
B_Q02bUK1_22			6	Valid skip	010
B_Q02bUK1_23	4	Education - Current qualification - Level - key Sk	-1	Missing	001
B_Q02bUK1_23			1	Marked	000
B_Q02bUK1_23			2	Not marked	100
B_Q02bUK1_23			6	Valid skip	010
B_Q02bUK1_24	4	Education - Current qualification - Level - Entry	-1	Missing	001
B_Q02bUK1_24			1	Marked	000
B_Q02bUK1_24			2	Not marked	100
B_Q02bUK1_24			6	Valid skip	010
B_Q02bUK1_25	4	Education - Current qualification - Level - Any ot	-1	Missing	001
B_Q02bUK1_25			1	Marked	000
B_Q02bUK1_25			2	Not marked	100
B_Q02bUK1_25			6	Valid skip	010
B_Q02bUK2	7	Education - Current qualification - NVQ/SVQ Level	-1	Missing	000001
B_Q02bUK2			1	Level 1	000000
B_Q02bUK2			2	Level 2	100000
B_Q02bUK2			3	Level 3	010000
B_Q02bUK2			4	Level 4	001000
B_Q02bUK2			5	Level 5	000100
B_Q02bUK2			6	Valid skip	000010
B_Q02bUK3	6	Education - Current qualification - BTEC/EdExcel/L	-1	Missing	00001
B_Q02bUK3			1	A higher Level (leve	00000
B_Q02bUK3			2	National Certificate	10000
B_Q02bUK3			3	First Diploma or gen	01000
B_Q02bUK3			4	First certificate or	00100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02bUK3			6	Valid skip	00010
B_Q02bUK4	7	Education - Current qualification - SCOTVEC Level	-1	Missing	000001
B_Q02bUK4			1	A higher Level (leve	000000
B_Q02bUK4			2	Full national certif	100000
B_Q02bUK4			3	A first diploma or g	010000
B_Q02bUK4			4	A first certificate	001000
B_Q02bUK4			5	Modules towards a Na	000100
B_Q02bUK4			6	Valid skip	000010
B_Q02bUK5	7	Education - Current qualification - National Quali	-1	Missing	000001
B_Q02bUK5			1	Access Level	000000
B_Q02bUK5			2	Intermediate 1	100000
B_Q02bUK5			3	Intermediate 2	010000
B_Q02bUK5			4	Higher	001000
B_Q02bUK5			5	Advanced Higher	000100
B_Q02bUK5			6	Valid skip	000010
B_Q02bUK6	6	Education - Current qualification - RSA Level	-1	Missing	00001
B_Q02bUK6			1	a higher diploma	00000
B_Q02bUK6			2	an advanced diploma	10000
B_Q02bUK6			3	a diploma	01000
B_Q02bUK6			4	or some other RSA (i	00100
B_Q02bUK6			6	Valid skip	00010
B_Q02bUK7	5	Education - Current qualification - City & Guilds	-1	Missing	0001
B_Q02bUK7			1	Advanced craft/part	0000
B_Q02bUK7			2	craft/part 2	1000
B_Q02bUK7			3	foundation/part 1	0100
B_Q02bUK7			6	Valid skip	0010
B_Q02bUK8	4	Education - Current qualification - Doing an appre	-1	Missing	001
B_Q02bUK8			1	Yes	000
B_Q02bUK8			2	No	100
B_Q02bUK8			6	Valid skip	010
B_Q02bUS	12	Education - Current qualification - Level	-1	Missing	00000000001
B_Q02bUS			1	Grades 1-6	00000000000
B_Q02bUS			2	Grades 7-9	10000000000
B_Q02bUS			3	High school diploma	01000000000
B_Q02bUS			4	Pre-associate educat	00100000000
B_Q02bUS			6	A certificate from a	00010000000
B_Q02bUS			7	Associate degree	00001000000
B_Q02bUS			8	Bachelor's degree (e	00000100000
B_Q02bUS			9	Master's degree (e.g	00000010000
B_Q02bUS			10	Professional degree	00000001000
B_Q02bUS			11	Doctorate degree (e.	00000000100
B_Q02bUS			96	Valid skip	00000000010
B_Q02cCZ	13	Education - Current qualification - Area of study	-1	Missing	000000000001
B_Q02cCZ			1	General programmes	000000000000
B_Q02cCZ			2	Teacher training and	100000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02cCZ			3	Humanities, language	010000000000
B_Q02cCZ			4	Social sciences	001000000000
B_Q02cCZ			5	Business and law	000100000000
B_Q02cCZ			6	Science, mathematics	000010000000
B_Q02cCZ			7	Engineering, manufac	000001000000
B_Q02cCZ			8	Agriculture and vete	000000100000
B_Q02cCZ			9	Health	000000010000
B_Q02cCZ			10	Welfare	000000001000
B_Q02cCZ			11	Services	000000000100
B_Q02cCZ			96	Valid skip	000000000010
B_Q02cKO	12	KO_Education - Current qualification - Area of stu	-1	Missing	00000000001
B_Q02cKO			1	General programmes	00000000000
B_Q02cKO			2	Teacher training and	10000000000
B_Q02cKO			3	Humanities, language	01000000000
B_Q02cKO			4	Social sciences, bus	00100000000
B_Q02cKO			5	Science, mathematics	00010000000
B_Q02cKO			6	Engineering, manufac	00001000000
B_Q02cKO			7	Agriculture and vete	00000100000
B_Q02cKO			8	Dental and medicine	00000010000
B_Q02cKO			9	Health and wellfare	00000001000
B_Q02cKO			10	Services	00000000100
B_Q02cKO			96	Valid skip	00000000010
B_Q02cNL	13	Education - Current qualification - Area of study	-1	Missing	000000000001
B_Q02cNL			1	general programmes	000000000000
B_Q02cNL			2	teacher training, ed	100000000000
B_Q02cNL			3	humanities, language	010000000000
B_Q02cNL			4	social sciences, com	001000000000
B_Q02cNL			5	economy, business, m	000100000000
B_Q02cNL			6	law, civil service,	000010000000
B_Q02cNL			7	mathematics, natural	000001000000
B_Q02cNL			8	technics	000000100000
B_Q02cNL			9	agriculture, veterin	000000010000
B_Q02cNL			10	health, welfare, per	000000001000
B_Q02cNL			11	tourism, horeca, tra	000000000100
B_Q02cNL			96	Valid skip	000000000010
B_Q02cUK	21	Education - Current qualification - Area of study	-1	Missing	00000000000000000001
B_Q02cUK			1	General programmes	00000000000000000000
B_Q02cUK			2	Medicine	10000000000000000000
B_Q02cUK			3	Medical related subj	01000000000000000000
B_Q02cUK			4	Biological Sciences	00100000000000000000
B_Q02cUK			5	Agricultural science	00010000000000000000
B_Q02cUK			6	Physical/Environment	00001000000000000000
B_Q02cUK			7	Mathematical Science	00000100000000000000
B_Q02cUK			8	Engineering	00000010000000000000
B_Q02cUK			9	Technology	00000001000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q02cUK			10	Architecture and rel	00000000100000000000
B_Q02cUK			11	Social Sciences (inc	00000000010000000000
B_Q02cUK			12	Business and Financi	00000000001000000000
B_Q02cUK			13	Librarianship and In	00000000000100000000
B_Q02cUK			14	Linguistics, English	00000000000010000000
B_Q02cUK			15	European Languages	00000000000001000000
B_Q02cUK			16	Other languages	00000000000000100000
B_Q02cUK			17	Humanities	00000000000000010000
B_Q02cUK			18	Arts	00000000000000001000
B_Q02cUK			19	Education	00000000000000000100
B_Q02cUK			96	Valid skip	00000000000000000010
B_Q03aAU	4	Education - Uncompleted qualification	-1	Missing	001
B_Q03aAU			1	Yes	000
B_Q03aAU			2	No	100
B_Q03aAU			6	Valid skip	010
B_Q03aDE	6	Education National - Uncompleted qualification	-1	Missing	00001
B_Q03aDE			1	Yes, school providin	00000
B_Q03aDE			2	Yes, professional tr	10000
B_Q03aDE			3	Yes, both of the abo	01000
B_Q03aDE			4	No	00100
B_Q03aDE			6	Valid skip	00010
B_Q03b3FR	15	Education - Uncompleted qualification - Level of f	-1	Missing	00000000000001
B_Q03b3FR			1	ISCED 1	00000000000000
B_Q03b3FR			2	ISCED 2	10000000000000
B_Q03b3FR			3	ISCED 3C shorter tha	01000000000000
B_Q03b3FR			4	ISCED 3C 2 years or	00100000000000
B_Q03b3FR			5	ISCED 3A-B	00010000000000
B_Q03b3FR			6	ISCED 3 (without dis	00001000000000
B_Q03b3FR			7	ISCED 4C	00000100000000
B_Q03b3FR			8	ISCED 4A-B	00000010000000
B_Q03b3FR			9	ISCED 4 (without dis	00000001000000
B_Q03b3FR			10	ISCED 5B	00000000100000
B_Q03b3FR			11	ISCED 5A, bachelor d	00000000010000
B_Q03b3FR			12	ISCED 5A, master deg	00000000001000
B_Q03b3FR			13	ISCED 6	00000000000100
B_Q03b3FR			96	Valid skip	00000000000010
B_Q03bAT	18	Education - Uncompleted qualification - Level - NA	-1	Missing	00000000000000001
B_Q03bAT			1	Compulsory school	00000000000000000
B_Q03bAT			2	Apprenticeship	10000000000000000
B_Q03bAT			3	Vocational School (<	01000000000000000
B_Q03bAT			4	Vocational School (2	00100000000000000
B_Q03bAT			5	Nursing	00010000000000000
B_Q03bAT			6	Master craftsman's c	00001000000000000
B_Q03bAT			7	Academic Secondary S	00000100000000000
B_Q03bAT			8	Vocational college	00000010000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q03bAT			9	Post-secondary cours	00000001000000000
B_Q03bAT			10	Post-secondary colle	00000000100000000
B_Q03bAT			11	University courses	00000000010000000
B_Q03bAT			12	University-Bachelor	00000000001000000
B_Q03bAT			13	University-Master	00000000000100000
B_Q03bAT			14	Post-graduate course	00000000000010000
B_Q03bAT			15	Doctoral Programme	00000000000001000
B_Q03bAT			16	Foreign Qualificatio	00000000000000100
B_Q03bAT			96	Valid skip	00000000000000010
B_Q03bAU	5	Education - Uncompleted qualification - Level not	-1	Missing	0001
B_Q03bAU			1	Level	0000
B_Q03bAU			2	Year 12 or equivalen	1000
B_Q03bAU			3	Statement of attainm	0100
B_Q03bAU			6	Valid skip	0010
B_Q03bBE	13	Education - Uncompleted qualification - Level	-1	Missing	000000000001
B_Q03bBE			1	ISCED 1	000000000000
B_Q03bBE			2	ISCED 2	100000000000
B_Q03bBE			3	ISCED 3C 2 years or	010000000000
B_Q03bBE			4	ISCED 3A-B	001000000000
B_Q03bBE			5	ISCED 3 (without dis	000100000000
B_Q03bBE			6	ISCED 4A-B	000010000000
B_Q03bBE			7	ISCED 5B	000001000000
B_Q03bBE			8	ISCED 5A, bachelor d	000000100000
B_Q03bBE			9	ISCED 5A, master deg	000000010000
B_Q03bBE			10	ISCED 6	000000001000
B_Q03bBE			11	Foreign qualificatio	000000000100
B_Q03bBE			96	Valid skip	000000000010
B_Q03bca1	16	Education - Uncompleted program of study - Level	-1	Missing	000000000000001
B_Q03bca1			1	Grade 6	000000000000000
B_Q03bca1			2	Less than high schoo	100000000000000
B_Q03bca1			3	High school diploma	010000000000000
B_Q03bca1			4	Trade/vocational cer	001000000000000
B_Q03bca1			5	Apprenticeship certi	000100000000000
B_Q03bca1			6	CEGEP diploma or cer	000010000000000
B_Q03bca1			7	Non-university certi	000001000000000
B_Q03bca1			8	University transfer	000000100000000
B_Q03bca1			9	University certifica	000000010000000
B_Q03bca1			10	Bachelor's degree	000000001000000
B_Q03bca1			11	University certifica	000000000100000
B_Q03bca1			12	First professional d	000000000010000
B_Q03bca1			13	Master's	000000000001000
B_Q03bca1			14	Ph.D.	000000000000100
B_Q03bca1			96	Valid skip	000000000000010
B_Q03bca2	4	Education - Uncompleted program of study - CEGEP d	-1	Missing	001
B_Q03bca2			1	Yes	000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q03bca2			2	No	100
B_Q03bca2			6	Valid skip	010
B_Q03bca3	9	Education - Uncompleted program of study - Length	-1	Missing	00000001
B_Q03bca3			1	Less than 3 months	00000000
B_Q03bca3			2	3 months to less tha	10000000
B_Q03bca3			3	One year	01000000
B_Q03bca3			4	Greater than one yea	00100000
B_Q03bca3			5	Two years	00010000
B_Q03bca3			6	Greater than two yea	00001000
B_Q03bca3			7	Three years or more	00000100
B_Q03bca3			96	Valid skip	00000010
B_Q03bCY	9	Education - Uncompleted qualification - Level	-1	Missing	00000001
B_Q03bCY			1	Primary school	00000000
B_Q03bCY			2	Public/Private Secon	10000000
B_Q03bCY			3	High School/Vocation	01000000
B_Q03bCY			4	Non-Univ. Degree/Dip	00100000
B_Q03bCY			5	Undergraduate degree	00010000
B_Q03bCY			6	Postgraduate degree,	00001000
B_Q03bCY			7	Doctorate	00000100
B_Q03bCY			96	Valid skip	00000010
B_Q03bCZ	14	Education - Uncompleted qualification - Level	-1	Missing	0000000000001
B_Q03bCZ			1	First level of basic	0000000000000
B_Q03bCZ			2	basic ISCED 2	1000000000000
B_Q03bCZ			3	vocational without m	0100000000000
B_Q03bCZ			4	vocational without m	0010000000000
B_Q03bCZ			5	ISCED 3A vocational	0001000000000
B_Q03bCZ			6	ISCED 3A technical w	0000100000000
B_Q03bCZ			7	ISCED 3A general wit	0000010000000
B_Q03bCZ			8	ISCED 4 follow-up co	0000001000000
B_Q03bCZ			9	ISCED 5B higher prof	0000000100000
B_Q03bCZ			10	ISCED 5A, bachelor	0000000010000
B_Q03bCZ			11	ISCED 5A, master	0000000001000
B_Q03bCZ			12	ISCED 6, post gradua	0000000000100
B_Q03bCZ			96	Valid skip	0000000000010
B_Q03bDE1	10	Education National - Uncompleted school qualificat	-1	Missing	000000001
B_Q03bDE1			1	Hauptschulabschluss	000000000
B_Q03bDE1			2	Realschulabschluss (	100000000
B_Q03bDE1			3	Polytechnische Obers	010000000
B_Q03bDE1			4	Polytechnische Obers	001000000
B_Q03bDE1			5	Fachhochschulereife,	000100000
B_Q03bDE1			6	Abitur/EOS (General	000010000
B_Q03bDE1			7	Abitur (General high	000001000
B_Q03bDE1			8	Another school leavi	000000100
B_Q03bDE1			96	Valid skip	000000010
B_Q03bDE2	12	Education National - Uncompleted professional qual	-1	Missing	00000000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q03bDE2			1	Completed Apprentice	00000000000
B_Q03bDE2			2	Basic vocational tra	10000000000
B_Q03bDE2			3	Training at Fachschu	01000000000
B_Q03bDE2			4	Berufsakademie, Fach	00100000000
B_Q03bDE2			5	Bachelor at Fachhoch	00010000000
B_Q03bDE2			6	Master/Diplom at Fac	00001000000
B_Q03bDE2			7	Bachelor at universi	00000100000
B_Q03bDE2			8	Master/Diplom at uni	00000010000
B_Q03bDE2			9	Doctorate	00000001000
B_Q03bDE2			10	Another professional	00000000100
B_Q03bDE2			96	Valid skip	00000000010
B_Q03bDK	15	Education - Uncompleted qualification - Level	-1	Missing	00000000000001
B_Q03bDK			1	Primary school, grad	00000000000000
B_Q03bDK			2	Lower secondary, gra	10000000000000
B_Q03bDK			3	Upper secondary voca	01000000000000
B_Q03bDK			4	Upper secondary voca	00100000000000
B_Q03bDK			5	Upper secondary gene	00010000000000
B_Q03bDK			6	Upper secondary unde	00001000000000
B_Q03bDK			7	Post secondary short	00000100000000
B_Q03bDK			8	Post secondary entra	00000010000000
B_Q03bDK			9	Post secondary non t	00000001000000
B_Q03bDK			10	Tertiary not researc	00000000100000
B_Q03bDK			11	Bachelor degree	00000000010000
B_Q03bDK			12	Master degree	00000000001000
B_Q03bDK			13	Ph.d or otther resea	00000000000100
B_Q03bDK			96	Valid skip	00000000000010
B_Q03bEE	19	Education - Uncompleted qualification - Level	-1	Missing	000000000000000001
B_Q03bEE			1	Primary education	000000000000000000
B_Q03bEE			2	Basic education	100000000000000000
B_Q03bEE			3	General secondary ed	010000000000000000
B_Q03bEE			4	Vocational education	001000000000000000
B_Q03bEE			5	Vocational education	000100000000000000
B_Q03bEE			6	Vocational education	000010000000000000
B_Q03bEE			7	Vocational secondary	000001000000000000
B_Q03bEE			8	Secondary specialise	000000100000000000
B_Q03bEE			9	Vocational secondary	000000010000000000
B_Q03bEE			10	Secondary specialise	000000001000000000
B_Q03bEE			11	Applied higher educa	000000000100000000
B_Q03bEE			12	Bachelor's degree (3	000000000010000000
B_Q03bEE			13	Bachelor's degree (4	000000000001000000
B_Q03bEE			14	Higher education (st	000000000000100000
B_Q03bEE			15	Master's degree (3+2	000000000000010000
B_Q03bEE			16	Master's degree (4+2	000000000000001000
B_Q03bEE			17	Doctoral degree (inc	000000000000000100
B_Q03bEE			96	Valid skip	000000000000000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q03bES	12	Education - Uncompleted qualification - Level	-1	Missing	00000000001
B_Q03bES			1	Not stated or inferr	00000000000
B_Q03bES			2	Not stated or inferr	10000000000
B_Q03bES			3	Not stated or inferr	01000000000
B_Q03bES			4	Not stated or inferr	00100000000
B_Q03bES			5	Bachillerato,. Y sim	00010000000
B_Q03bES			6	Pruebas de acceso a	00001000000
B_Q03bES			7	Pruebas de acceso a	00000100000
B_Q03bES			8	Pruebas de acceso a	00000010000
B_Q03bES			9	Pruebas de aster y e	00000001000
B_Q03bES			10	Programas de doctora	00000000100
B_Q03bES			96	Valid skip	00000000010
B_Q03bFI	12	Education - Uncompleted qualification - Level	-1	Missing	00000000001
B_Q03bFI			1	ISCED 1	00000000000
B_Q03bFI			2	ISCED 2	10000000000
B_Q03bFI			3	Upper secondary voca	01000000000
B_Q03bFI			4	General upper second	00100000000
B_Q03bFI			5	Specialist vocationa	00010000000
B_Q03bFI			6	Vocational post-seco	00001000000
B_Q03bFI			7	Polytechnic degree (	00000100000
B_Q03bFI			8	Bachelor's degree (I	00000010000
B_Q03bFI			9	Master's degree (ISC	00000001000
B_Q03bFI			10	Licentiate's and doc	00000000100
B_Q03bFI			96	Valid skip	00000000010
B_Q03bFR1	16	Education - Uncompleted qualification - Level	-1	Missing	000000000000001
B_Q03bFR1			1	ISCED 1	000000000000000
B_Q03bFR1			2	ISCED 2	100000000000000
B_Q03bFR1			3	ISCED 3C shorter tha	010000000000000
B_Q03bFR1			4	ISCED 3C 2 years or	001000000000000
B_Q03bFR1			5	ISCED 3A-B	000100000000000
B_Q03bFR1			6	ISCED 3 (without dis	000010000000000
B_Q03bFR1			7	ISCED 4C	000001000000000
B_Q03bFR1			8	ISCED 4A-B	000000100000000
B_Q03bFR1			9	ISCED 4 (without dis	000000010000000
B_Q03bFR1			10	ISCED 5B	000000001000000
B_Q03bFR1			11	ISCED 5A, bachelor d	000000000100000
B_Q03bFR1			12	ISCED 5A, master deg	000000000010000
B_Q03bFR1			13	ISCED 6	000000000001000
B_Q03bFR1			14	Foreign qualificatio	000000000000100
B_Q03bFR1			96	Valid skip	000000000000010
B_Q03bIE	15	Education - Uncompleted qualification - Level	-1	Missing	00000000000001
B_Q03bIE			1	No formal education	00000000000000
B_Q03bIE			2	Primary education (o	10000000000000
B_Q03bIE			3	Secondary 1 (Junior/	01000000000000
B_Q03bIE			4	Transition year prog	00100000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q03bIE			5	Secondary 2 (Leaving	00010000000000
B_Q03bIE			6	Technical or Vocatio	00001000000000
B_Q03bIE			7	Advanced Certificate	00000100000000
B_Q03bIE			8	Higher Certificate (	00000010000000
B_Q03bIE			9	Diploma (e.g. Nation	00000001000000
B_Q03bIE			10	Honours Bachelor Deg	00000000100000
B_Q03bIE			11	Professional (Honour	00000000010000
B_Q03bIE			12	Post-Graduate (e.g.	00000000001000
B_Q03bIE			13	Doctorate or higher	00000000000100
B_Q03bIE			96	Valid skip	00000000000010
B_Q03bIT	12	Education - Uncompleted qualification - Level	-1	Missing	00000000001
B_Q03bIT			1	Primary education or	00000000000
B_Q03bIT			2	Lower secondary or s	10000000000
B_Q03bIT			3	Regional Vocational	01000000000
B_Q03bIT			4	Educational and voca	00100000000
B_Q03bIT			5	Upper secondary educ	00010000000
B_Q03bIT			6	Post-second. non ter	00001000000
B_Q03bIT			7	Music Conservatory D	00000100000
B_Q03bIT			8	First stage of terti	00000010000
B_Q03bIT			9	First or second leve	00000001000
B_Q03bIT			10	Research Doctoral de	00000000100
B_Q03bIT			96	Valid skip	00000000010
B_Q03bJP	14	Education - Uncompleted qualification - Level	-1	Missing	0000000000001
B_Q03bJP			1	Elementary school	0000000000000
B_Q03bJP			2	Lower secondary scho	1000000000000
B_Q03bJP			3	Short-term course of	0100000000000
B_Q03bJP			4	Specialized course o	0010000000000
B_Q03bJP			5	General/integrated c	0001000000000
B_Q03bJP			6	Passed upper seconda	0000100000000
B_Q03bJP			7	Advanced course of u	0000010000000
B_Q03bJP			8	Regular/advanced cou	0000001000000
B_Q03bJP			9	Undergraduate progra	0000000100000
B_Q03bJP			10	Master's program/Doc	0000000010000
B_Q03bJP			11	Doctoral programs of	0000000001000
B_Q03bJP			12	Specialized training	0000000000100
B_Q03bJP			96	Valid skip	0000000000010
B_Q03bKO	12	KO_Education - Uncompleted qualification - Level	-1	Missing	00000000001
B_Q03bKO			1	Elementary school	00000000000
B_Q03bKO			2	Middle school	10000000000
B_Q03bKO			3	High school(college	01000000000
B_Q03bKO			4	High school(vocation	00100000000
B_Q03bKO			5	2-3 year college	00010000000
B_Q03bKO			6	4 year college(speci	00001000000
B_Q03bKO			7	4 year college(gener	00000100000
B_Q03bKO			8	Master's degree(spec	00000010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q03bKO			9	Master's degree(gene	00000001000
B_Q03bKO			10	Doctoral degree	00000000100
B_Q03bKO			96	Valid skip	00000000010
B_Q03bNL	17	Education - Uncompleted qualification - Level	-1	Missing	0000000000000001
B_Q03bNL			1	primary education (i	0000000000000000
B_Q03bNL			2	sec education, first	1000000000000000
B_Q03bNL			3	sec education, first	0100000000000000
B_Q03bNL			4	secondary education,	0010000000000000
B_Q03bNL			5	secondary education,	0001000000000000
B_Q03bNL			6	secondary education,	0000100000000000
B_Q03bNL			7	secondary education,	0000010000000000
B_Q03bNL			8	secondary education,	0000001000000000
B_Q03bNL			9	sec education, secon	0000000100000000
B_Q03bNL			10	secondary education,	0000000010000000
B_Q03bNL			11	tertiary education,	0000000001000000
B_Q03bNL			12	tertiary education,	0000000000100000
B_Q03bNL			13	tertiary education,	0000000000010000
B_Q03bNL			14	tertiary education,	0000000000001000
B_Q03bNL			15	tertiary education,	0000000000000100
B_Q03bNL			96	Valid skip	0000000000000010
B_Q03bNO	13	Education - Uncompleted qualification - Level	-1	Missing	000000000001
B_Q03bNO			1	ISCED 1	000000000000
B_Q03bNO			2	ISCED 2	100000000000
B_Q03bNO			3	ISCED 3C shorter tha	010000000000
B_Q03bNO			4	ISCED 3C 2 years or	001000000000
B_Q03bNO			5	ISCED 3A-B	000100000000
B_Q03bNO			6	ISCED 4C	000010000000
B_Q03bNO			7	ISCED 4A-B	000001000000
B_Q03bNO			8	ISCED 5B	000000100000
B_Q03bNO			9	ISCED 5A, bachelor d	000000010000
B_Q03bNO			10	ISCED 5A, Master deg	000000001000
B_Q03bNO			11	ISCED 6	000000000100
B_Q03bNO			96	Valid skip	000000000010
B_Q03bPL	11	Education - Uncompleted qualification - Level	-1	Missing	0000000001
B_Q03bPL			1	ISCED 1	0000000000
B_Q03bPL			2	ISCED 2	1000000000
B_Q03bPL			3	ISCED 3C	0100000000
B_Q03bPL			4	ISCED 3B	0010000000
B_Q03bPL			5	ISCED 3A	0001000000
B_Q03bPL			6	ISCED 4	0000100000
B_Q03bPL			7	BA, ISCED 5A (I degr	0000010000
B_Q03bPL			8	MA, ISCED 5A (II deg	0000001000
B_Q03bPL			9	ISCED 6	0000000100
B_Q03bPL			96	Valid skip	0000000010
B_Q03bRU	10	Education - Uncompleted qualification - Level	-1	Missing	000000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q03bRU			1	ISCED 1	000000000
B_Q03bRU			2	ISCED 2	100000000
B_Q03bRU			3	ISCED 3 (without dis	010000000
B_Q03bRU			4	ISCED 4 (without dis	001000000
B_Q03bRU			5	ISCED 5B	000100000
B_Q03bRU			6	ISCED 5A, bachelor d	000010000
B_Q03bRU			7	ISCED 5A, master deg	000001000
B_Q03bRU			8	ISCED 6	000000100
B_Q03bRU			96	Valid skip	000000010
B_Q03bSE	18	Education - Uncompleted qualification - Level	-1	Missing	00000000000000001
B_Q03bSE			1	Not stated or inferr	00000000000000000
B_Q03bSE			2	Not stated or inr	10000000000000000
B_Q03bSE			3	Grundskola, enhetssk	01000000000000000
B_Q03bSE			4	Yrkesutbildning	00100000000000000
B_Q03bSE			5	Grundskolekompetens	00010000000000000
B_Q03bSE			6	Flickskola	00001000000000000
B_Q03bSE			7	Gymnasie fackskola y	00000100000000000
B_Q03bSE			8	Gymnasie fackskola y	00000010000000000
B_Q03bSE			9	Gymnasie fackskola y	00000001000000000
B_Q03bSE			10	Vuxenutbildning mots	00000000100000000
B_Q03bSE			11	Vuxenutbildning mots	00000000010000000
B_Q03bSE			12	Eftergymnasial utbil	00000000001000000
B_Q03bSE			13	Eftergymnasial utbil	00000000000100000
B_Q03bSE			14	Eftergymnasial utbil	00000000000010000
B_Q03bSE			15	Eftergymnasial utbil	00000000000001000
B_Q03bSE			16	Forskarutbildning	00000000000000100
B_Q03bSE			96	Valid skip	00000000000000010
B_Q03bSK	12	Education - Uncompleted qualification - Level	-1	Missing	00000000001
B_Q03bSK			1	Primary school 1-4.	00000000000
B_Q03bSK			2	Primary school 5.-9.	10000000000
B_Q03bSK			3	Secondary technical	01000000000
B_Q03bSK			4	Secondary technical	00100000000
B_Q03bSK			5	Secondary schools wi	00010000000
B_Q03bSK			6	Upper secondary scho	00001000000
B_Q03bSK			7	Pre-tertiary school,	00000100000
B_Q03bSK			8	Bachelor degree, Gra	00000010000
B_Q03bSK			9	Master degree	00000001000
B_Q03bSK			10	PhD studies, Second	00000000100
B_Q03bSK			96	Valid skip	00000000010
B_Q03bUK1	28	Education - Uncompleted qualification - Level	-1	Missing	000000000000000000000000001
B_Q03bUK1			1	Degree level qualifi	000000000000000000000000000
B_Q03bUK1			2	Diploma in higher ed	100000000000000000000000000
B_Q03bUK1			3	HNC/HND	010000000000000000000000000
B_Q03bUK1			4	ONC/OND	001000000000000000000000000
B_Q03bUK1			5	BTEC, BEC, TEC or Ed	000100000000000000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q03bUK1			6	SCOTVEC, SCOTEC, SCO	000010000000000000000000000
B_Q03bUK1			7	Teaching qualificati	000001000000000000000000000
B_Q03bUK1			8	Nursing or other med	000000100000000000000000000
B_Q03bUK1			9	other Higher Educati	000000010000000000000000000
B_Q03bUK1			10	A Level/Vocational A	000000001000000000000000000
B_Q03bUK1			11	Highers (Scotland)	000000000100000000000000000
B_Q03bUK1			12	NVQ/SVQ	000000000010000000000000000
B_Q03bUK1			13	GNVQ/GSVQ	000000000001000000000000000
B_Q03bUK1			14	AS Level/Vocational	000000000000100000000000000
B_Q03bUK1			15	Advanced highers or	000000000000010000000000000
B_Q03bUK1			16	Access to HE	000000000000001000000000000
B_Q03bUK1			17	O Level/GCSE/Vocatio	000000000000000100000000000
B_Q03bUK1			18	Intermediate 1 or 2	000000000000000010000000000
B_Q03bUK1			19	Standard Grade or O	000000000000000001000000000
B_Q03bUK1			20	National Qualificati	000000000000000000100000000
B_Q03bUK1			21	RSA/OCR	000000000000000000010000000
B_Q03bUK1			22	City and Guilds	000000000000000000001000000
B_Q03bUK1			23	YT Certificate/YTP	000000000000000000000100000
B_Q03bUK1			24	Key skills/Basic ski	000000000000000000000010000
B_Q03bUK1			25	Entry level qualific	000000000000000000000001000
B_Q03bUK1			26	Any other profession	000000000000000000000000100
B_Q03bUK1			96	Valid skip	000000000000000000000000010
B_Q03bUK2	7	Education - Uncompleted qualification - NVQ/SVQ Le	-1	Missing	000001
B_Q03bUK2			1	Level 1	000000
B_Q03bUK2			2	Level 2	100000
B_Q03bUK2			3	Level 3	010000
B_Q03bUK2			4	Level 4	001000
B_Q03bUK2			5	Level 5	000100
B_Q03bUK2			6	Valid skip	000010
B_Q03bUK3	6	Education - Uncompleted qualification - BTEC/BEC/T	-1	Missing	00001
B_Q03bUK3			1	A higher Level (leve	00000
B_Q03bUK3			2	National Certificate	10000
B_Q03bUK3			3	First Diploma or gen	01000
B_Q03bUK3			4	First certificate or	00100
B_Q03bUK3			6	Valid skip	00010
B_Q03bUK4	7	Education - Uncompleted qualification - SCOTVEC/SC	-1	Missing	000001
B_Q03bUK4			1	A higher Level (leve	000000
B_Q03bUK4			2	Full national certif	100000
B_Q03bUK4			3	A first diploma or g	010000
B_Q03bUK4			4	A first certificate	001000
B_Q03bUK4			5	Modules towards a Na	000100
B_Q03bUK4			6	Valid skip	000010
B_Q03bUK5	7	Education - Uncompleted qualification - National Q	-1	Missing	000001
B_Q03bUK5			1	Access Level	000000
B_Q03bUK5			2	Intermediate 1	100000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q03bUK5			3	Intermediate 2	010000
B_Q03bUK5			4	Higher	001000
B_Q03bUK5			5	Advanced Higher	000100
B_Q03bUK5			6	Valid skip	000010
B_Q03bUK6	6	Education - Uncompleted qualification - RSA Level	-1	Missing	00001
B_Q03bUK6			1	a higher diploma	00000
B_Q03bUK6			2	an advanced diploma	10000
B_Q03bUK6			3	a diploma	01000
B_Q03bUK6			4	or some other RSA (i	00100
B_Q03bUK6			6	Valid skip	00010
B_Q03bUK7	5	Education - Uncompleted qualification - City & Gui	-1	Missing	0001
B_Q03bUK7			1	Advanced craft/part	0000
B_Q03bUK7			2	craft/part 2	1000
B_Q03bUK7			3	foundation/part 1	0100
B_Q03bUK7			6	Valid skip	0010
B_Q03bUK8	4	Education - Uncompleted qualification - doing Appr	-1	Missing	001
B_Q03bUK8			1	Yes	000
B_Q03bUK8			2	No	100
B_Q03bUK8			6	Valid skip	010
B_Q03bUS	12	Education - Uncompleted qualification - Level	-1	Missing	00000000001
B_Q03bUS			1	Grades 1-6	00000000000
B_Q03bUS			2	Grades 7-9	10000000000
B_Q03bUS			3	High school diploma	01000000000
B_Q03bUS			4	Pre-associate educat	00100000000
B_Q03bUS			6	A certificate from a	00010000000
B_Q03bUS			7	Associate degree	00001000000
B_Q03bUS			8	Bachelor's degree (e	00000100000
B_Q03bUS			9	Master's degree (e.g	00000010000
B_Q03bUS			10	Professional degree	00000001000
B_Q03bUS			11	Doctorate degree (e.	00000000100
B_Q03bUS			96	Valid skip	00000000010
B_Q04aAU	4	Education - Formal qualification - Last 12 months	-1	Missing	001
B_Q04aAU			1	Yes	000
B_Q04aAU			2	No	100
B_Q04aAU			6	Valid skip	010
B_Q04aDE	6	Education National - Last 12 months	-1	Missing	00001
B_Q04aDE			1	Yes, school providin	00000
B_Q04aDE			2	Yes, professional tr	10000
B_Q04aDE			3	Yes, both of the abo	01000
B_Q04aDE			4	No	00100
B_Q04aDE			6	Valid skip	00010
B_Q04bDE	4	Education National - Type of qualification last at	-1	Missing	001
B_Q04bDE			1	General education qu	000
B_Q04bDE			2	Professional trainin	100
B_Q04bDE			6	Valid skip	010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q05a2RU	9	Education - Formal qualification - Country of fore	-1	Missing	00000001
B_Q05a2RU			1	Country 1	00000000
B_Q05a2RU			2	Country 2	10000000
B_Q05a2RU			3	Country 3	01000000
B_Q05a2RU			4	Country 4	00100000
B_Q05a2RU			5	Country 5	00010000
B_Q05a2RU			6	Country 6	00001000
B_Q05a2RU			7	Other country	00000100
B_Q05a2RU			96	Valid skip	00000010
B_Q05aAT	20	Education - Formal qualification - Level - NATIONA	-1	Missing	0000000000000000001
B_Q05aAT			1	Lower secondary Scho	0000000000000000000
B_Q05aAT			2	Prevocational School	1000000000000000000
B_Q05aAT			3	Apprenticeship	0100000000000000000
B_Q05aAT			4	Vocational School (<	0010000000000000000
B_Q05aAT			5	Vocational School (2	0001000000000000000
B_Q05aAT			6	Nursing	0000100000000000000
B_Q05aAT			7	Master craftsman's c	0000010000000000000
B_Q05aAT			8	Academic secondary s	0000001000000000000
B_Q05aAT			9	1-3rd Class in a Voc	0000000100000000000
B_Q05aAT			10	4 or 5th Class in a	0000000010000000000
B_Q05aAT			11	Post-secondary cours	0000000001000000000
B_Q05aAT			12	Post-secondary colle	0000000000100000000
B_Q05aAT			13	University courses	0000000000010000000
B_Q05aAT			14	University-Bachelor	0000000000001000000
B_Q05aAT			15	University-Master	0000000000000100000
B_Q05aAT			16	Post-graduate course	0000000000000010000
B_Q05aAT			17	Doctoral Programme	0000000000000001000
B_Q05aAT			18	Foreign Qualificatio	0000000000000000100
B_Q05aAT			96	Valid skip	0000000000000000010
B_Q05aAU	4	Education - Formal qualification - Last 12 months	-1	Missing	001
B_Q05aAU			1	Level	000
B_Q05aAU			2	Statement of attainm	100
B_Q05aAU			6	Valid skip	010
B_Q05aBE	13	Education - Formal qualification - Level	-1	Missing	000000000001
B_Q05aBE			1	ISCED 1	000000000000
B_Q05aBE			2	ISCED 2	100000000000
B_Q05aBE			3	ISCED 3C 2 years or	010000000000
B_Q05aBE			4	ISCED 3A-B	001000000000
B_Q05aBE			5	ISCED 3 (without dis	000100000000
B_Q05aBE			6	ISCED 4A-B	000010000000
B_Q05aBE			7	ISCED 5B	000001000000
B_Q05aBE			8	ISCED 5A, bachelor d	000000100000
B_Q05aBE			9	ISCED 5A, master deg	000000010000
B_Q05aBE			10	ISCED 6	000000001000
B_Q05aBE			11	Foreign qualificatio	000000000100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q05aBE			96	Valid skip	000000000010
B_Q05aca1	16	Education - Formal education - Level	-1	Missing	000000000000001
B_Q05aca1			1	Grade 6	000000000000000
B_Q05aca1			2	Less than high schoo	100000000000000
B_Q05aca1			3	High school diploma	010000000000000
B_Q05aca1			4	Trade/vocational cer	001000000000000
B_Q05aca1			5	Apprenticeship certi	000100000000000
B_Q05aca1			6	CEGEP diploma or cer	000010000000000
B_Q05aca1			7	Non-university certi	000001000000000
B_Q05aca1			8	University transfer	000000100000000
B_Q05aca1			9	University certifica	000000010000000
B_Q05aca1			10	Bachelor's degree	000000001000000
B_Q05aca1			11	University certifica	000000000100000
B_Q05aca1			12	First professional d	000000000010000
B_Q05aca1			13	Master's	000000000001000
B_Q05aca1			14	Ph.D.	000000000000100
B_Q05aca1			96	Valid skip	000000000000010
B_Q05aca2	4	Education - Formal education - CEGEP diploma/certi	-1	Missing	001
B_Q05aca2			1	Yes	000
B_Q05aca2			2	No	100
B_Q05aca2			6	Valid skip	010
B_Q05aca3	9	Education - Formal education - Length - Complete t	-1	Missing	00000001
B_Q05aca3			1	Less than 3 months	00000000
B_Q05aca3			2	3 months to less tha	10000000
B_Q05aca3			3	One year	01000000
B_Q05aca3			4	Greater than one yea	00100000
B_Q05aca3			5	Two years	00010000
B_Q05aca3			6	Greater than two yea	00001000
B_Q05aca3			7	Three years or more	00000100
B_Q05aca3			96	Valid skip	00000010
B_Q05aCY	9	Education - Formal qualification - Level	-1	Missing	00000001
B_Q05aCY			1	Primary school	00000000
B_Q05aCY			2	Public/Private Secon	10000000
B_Q05aCY			3	High School/Vocation	01000000
B_Q05aCY			4	Non-Univ. Degree/Dip	00100000
B_Q05aCY			5	Undergraduate degree	00010000
B_Q05aCY			6	Postgraduate degree,	00001000
B_Q05aCY			7	Doctorate	00000100
B_Q05aCY			96	Valid skip	00000010
B_Q05aCZ	14	Education - Formal qualification - Level	-1	Missing	0000000000001
B_Q05aCZ			1	First level of basic	0000000000000
B_Q05aCZ			2	basic ISCED 2	1000000000000
B_Q05aCZ			3	vocational without m	0100000000000
B_Q05aCZ			4	vocational without m	0010000000000
B_Q05aCZ			5	ISCED 3A vocational	0001000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q05aCZ			6	ISCED 3A technical w	0000100000000
B_Q05aCZ			7	ISCED 3A general wit	0000010000000
B_Q05aCZ			8	ISCED 4 follow-up co	0000001000000
B_Q05aCZ			9	ISCED 5B higher prof	0000000100000
B_Q05aCZ			10	ISCED 5A, bachelor	0000000010000
B_Q05aCZ			11	ISCED 5A, master	0000000001000
B_Q05aCZ			12	ISCED 6, post gradua	0000000000100
B_Q05aCZ			96	Valid skip	0000000000010
B_Q05aDE1	8	Education National - Formal school qualification -	-1	Missing	0000001
B_Q05aDE1			1	Hauptschulabschluss	0000000
B_Q05aDE1			2	Realschulabschluss (	1000000
B_Q05aDE1			3	Fachhochschulreife,	0100000
B_Q05aDE1			4	Abitur/EOS (General	0010000
B_Q05aDE1			5	Abitur (General high	0001000
B_Q05aDE1			6	Another school leavi	0000100
B_Q05aDE1			96	Valid skip	0000010
B_Q05aDE2	12	Education National - Formal professional qualifica	-1	Missing	00000000001
B_Q05aDE2			1	Completed Apprentice	00000000000
B_Q05aDE2			2	Basic vocational tra	10000000000
B_Q05aDE2			3	Training at Fachschu	01000000000
B_Q05aDE2			4	Berufsakademie, Fach	00100000000
B_Q05aDE2			5	Bachelor at Fachhoch	00010000000
B_Q05aDE2			6	Master/Diplom at Fac	00001000000
B_Q05aDE2			7	Bachelor at universi	00000100000
B_Q05aDE2			8	Master/Diplom at uni	00000010000
B_Q05aDE2			9	Doctorate	00000001000
B_Q05aDE2			10	Another professional	00000000100
B_Q05aDE2			96	Valid skip	00000000010
B_Q05aDK	15	What was the level of this qualification?	-1	Missing	00000000000001
B_Q05aDK			1	Primary school, grad	00000000000000
B_Q05aDK			2	Lower secondary, gra	10000000000000
B_Q05aDK			3	Upper secondary voca	01000000000000
B_Q05aDK			4	Upper secondary voca	00100000000000
B_Q05aDK			5	Upper secondary gene	00010000000000
B_Q05aDK			6	Upper secondary unde	00001000000000
B_Q05aDK			7	Post secondary short	00000100000000
B_Q05aDK			8	Post secondary entra	00000010000000
B_Q05aDK			9	Post secondary non t	00000001000000
B_Q05aDK			10	Tertiary not researc	00000000100000
B_Q05aDK			11	Bachelor degree	00000000010000
B_Q05aDK			12	Master degree	00000000001000
B_Q05aDK			13	Ph.d or otther resea	00000000000100
B_Q05aDK			96	Valid skip	00000000000010
B_Q05aEE	14	Education - Formal qualification - Level	-1	Missing	0000000000001
B_Q05aEE			1	Primary education (1	0000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q05aEE			2	Basic education (7-9	1000000000000
B_Q05aEE			3	General secondary ed	0100000000000
B_Q05aEE			4	Vocational education	0010000000000
B_Q05aEE			5	Vocational education	0001000000000
B_Q05aEE			6	Voc ed on the basis	0000100000000
B_Q05aEE			7	Vocational secondary	0000010000000
B_Q05aEE			8	Vocational secondary	0000001000000
B_Q05aEE			9	Applied higher educa	0000000100000
B_Q05aEE			10	Bachelor's degree (3	0000000010000
B_Q05aEE			11	Master's degree (3+2	0000000001000
B_Q05aEE			12	Doctoral degree	0000000000100
B_Q05aEE			96	Valid skip	0000000000010
B_Q05aES	12	Education - Formal qualification - Level	-1	Missing	00000000001
B_Q05aES			1	Not stated or inferr	00000000000
B_Q05aES			2	Not stated or inferr	10000000000
B_Q05aES			3	Not stated or inferr	01000000000
B_Q05aES			4	Not stated or inferr	00100000000
B_Q05aES			5	Bachillerato,. Y sim	00010000000
B_Q05aES			6	Pruebas de acceso a	00001000000
B_Q05aES			7	Pruebas de acceso a	00000100000
B_Q05aES			8	Pruebas de acceso a	00000010000
B_Q05aES			9	Pruebas de aster y e	00000001000
B_Q05aES			10	Programas de doctora	00000000100
B_Q05aES			96	Valid skip	00000000010
B_Q05aFI	12	Education - Formal qualification - Level	-1	Missing	00000000001
B_Q05aFI			1	ISCED 1	00000000000
B_Q05aFI			2	ISCED 2	10000000000
B_Q05aFI			3	Upper secondary voca	01000000000
B_Q05aFI			4	General upper second	00100000000
B_Q05aFI			5	Specialist vocationa	00010000000
B_Q05aFI			6	Vocational post-seco	00001000000
B_Q05aFI			7	Polytechnic degree (	00000100000
B_Q05aFI			8	Bachelor's degree (I	00000010000
B_Q05aFI			9	Master's degree (ISC	00000001000
B_Q05aFI			10	Licentiate's and doc	00000000100
B_Q05aFI			96	Valid skip	00000000010
B_Q05aFR1	16	Education - Formal qualification - Level	-1	Missing	000000000000001
B_Q05aFR1			1	ISCED 1	000000000000000
B_Q05aFR1			2	ISCED 2	100000000000000
B_Q05aFR1			3	ISCED 3C shorter tha	010000000000000
B_Q05aFR1			4	ISCED 3C 2 years or	001000000000000
B_Q05aFR1			5	ISCED 3A-B	000100000000000
B_Q05aFR1			6	ISCED 3 (without dis	000010000000000
B_Q05aFR1			7	ISCED 4C	000001000000000
B_Q05aFR1			8	ISCED 4A-B	000000100000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q05aFR1			9	ISCED 4 (without dis	000000010000000
B_Q05aFR1			10	ISCED 5B	000000001000000
B_Q05aFR1			11	ISCED 5A, bachelor d	000000000100000
B_Q05aFR1			12	ISCED 5A, master deg	000000000010000
B_Q05aFR1			13	ISCED 6	000000000001000
B_Q05aFR1			14	Foreign qualificatio	000000000000100
B_Q05aFR1			96	Valid skip	000000000000010
B_Q05aIE	15	Education - Formal qualification - Level	-1	Missing	00000000000001
B_Q05aIE			1	No formal education	00000000000000
B_Q05aIE			2	Primary education (o	10000000000000
B_Q05aIE			3	Secondary 1 (Junior/	01000000000000
B_Q05aIE			4	Transition year prog	00100000000000
B_Q05aIE			5	Secondary 2 (Leaving	00010000000000
B_Q05aIE			6	Technical or Vocatio	00001000000000
B_Q05aIE			7	Advanced Certificate	00000100000000
B_Q05aIE			8	Higher Certificate (	00000010000000
B_Q05aIE			9	Diploma (e.g. Nation	00000001000000
B_Q05aIE			10	Honours Bachelor Deg	00000000100000
B_Q05aIE			11	Professional (Honour	00000000010000
B_Q05aIE			12	Post-Graduate (e.g.	00000000001000
B_Q05aIE			13	Doctorate or higher	00000000000100
B_Q05aIE			96	Valid skip	00000000000010
B_Q05aIT	12	Education - Formal qualification - Level	-1	Missing	00000000001
B_Q05aIT			1	Primary education or	00000000000
B_Q05aIT			2	Lower secondary or s	10000000000
B_Q05aIT			3	Regional Vocational	01000000000
B_Q05aIT			4	Educational and voca	00100000000
B_Q05aIT			5	Upper secondary educ	00010000000
B_Q05aIT			6	Post-second. non ter	00001000000
B_Q05aIT			7	Music Conservatory D	00000100000
B_Q05aIT			8	First stage of terti	00000010000
B_Q05aIT			9	First or second leve	00000001000
B_Q05aIT			10	Research Doctoral de	00000000100
B_Q05aIT			96	Valid skip	00000000010
B_Q05aJP	14	Education - Formal qualification - Level	-1	Missing	0000000000001
B_Q05aJP			1	Elementary school	0000000000000
B_Q05aJP			2	Lower secondary scho	1000000000000
B_Q05aJP			3	Short-term course of	0100000000000
B_Q05aJP			4	Specialized course o	0010000000000
B_Q05aJP			5	General/integrated c	0001000000000
B_Q05aJP			6	Passed upper seconda	0000100000000
B_Q05aJP			7	Advanced course of u	0000010000000
B_Q05aJP			8	Regular/advanced cou	0000001000000
B_Q05aJP			9	Undergraduate progra	0000000100000
B_Q05aJP			10	Master's program/Doc	0000000010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q05aJP			11	Doctoral programs of	0000000001000
B_Q05aJP			12	Specialized training	0000000000100
B_Q05aJP			96	Valid skip	0000000000010
B_Q05aKO	12	KO_Education - Formal qualification - Level	-1	Missing	00000000001
B_Q05aKO			1	Elementary school	00000000000
B_Q05aKO			2	Middle school	10000000000
B_Q05aKO			3	High school(college	01000000000
B_Q05aKO			4	High school(vocation	00100000000
B_Q05aKO			5	2-3 year college	00010000000
B_Q05aKO			6	4 year college(speci	00001000000
B_Q05aKO			7	4 year college(gener	00000100000
B_Q05aKO			8	Master's degree(spec	00000010000
B_Q05aKO			9	Master's degree(gene	00000001000
B_Q05aKO			10	Doctoral degree	00000000100
B_Q05aKO			96	Valid skip	00000000010
B_Q05aNL	17	Education - Formal qualification - Level	-1	Missing	0000000000000001
B_Q05aNL			1	primary education (i	0000000000000000
B_Q05aNL			2	sec education,first	1000000000000000
B_Q05aNL			3	sec education, first	0100000000000000
B_Q05aNL			4	secondary education,	0010000000000000
B_Q05aNL			5	secondary education,	0001000000000000
B_Q05aNL			6	secondary education,	0000100000000000
B_Q05aNL			7	secondary education,	0000010000000000
B_Q05aNL			8	secondary education,	0000001000000000
B_Q05aNL			9	sec education, secon	0000000100000000
B_Q05aNL			10	secondary education,	0000000010000000
B_Q05aNL			11	tertiary education,	0000000001000000
B_Q05aNL			12	tertiary education,	0000000000100000
B_Q05aNL			13	tertiary education,	0000000000010000
B_Q05aNL			14	tertiary education,	0000000000001000
B_Q05aNL			15	tertiary education,	0000000000000100
B_Q05aNL			96	Valid skip	0000000000000010
B_Q05aNO	13	Education - Formal qualification - Level	-1	Missing	000000000001
B_Q05aNO			1	ISCED 1	000000000000
B_Q05aNO			2	ISCED 2	100000000000
B_Q05aNO			3	ISCED 3C shorter tha	010000000000
B_Q05aNO			4	ISCED 3C 2 years or	001000000000
B_Q05aNO			5	ISCED 3A-B	000100000000
B_Q05aNO			6	ISCED 4C	000010000000
B_Q05aNO			7	ISCED 4A-B	000001000000
B_Q05aNO			8	ISCED 5B	000000100000
B_Q05aNO			9	ISCED 5A, bachelor d	000000010000
B_Q05aNO			10	ISCED 5A, Master deg	000000001000
B_Q05aNO			11	ISCED 6	000000000100
B_Q05aNO			96	Valid skip	000000000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q05aPL	11	Education - Formal qualification - Level	-1	Missing	0000000001
B_Q05aPL			1	ISCED 1	0000000000
B_Q05aPL			2	ISCED 2	1000000000
B_Q05aPL			3	ISCED 3C	0100000000
B_Q05aPL			4	ISCED 3B	0010000000
B_Q05aPL			5	ISCED 3A	0001000000
B_Q05aPL			6	ISCED 4	0000100000
B_Q05aPL			7	BA, ISCED 5A (I degr	0000010000
B_Q05aPL			8	MA, ISCED 5A (II deg	0000001000
B_Q05aPL			9	ISCED 6	0000000100
B_Q05aPL			96	Valid skip	0000000010
B_Q05aRU	10	Education - Formal qualification - Level	-1	Missing	000000001
B_Q05aRU			1	ISCED 1	000000000
B_Q05aRU			2	ISCED 2	100000000
B_Q05aRU			3	ISCED 3 (without dis	010000000
B_Q05aRU			4	ISCED 4 (without dis	001000000
B_Q05aRU			5	ISCED 5B	000100000
B_Q05aRU			6	ISCED 5A, bachelor d	000010000
B_Q05aRU			7	ISCED 5A, master deg	000001000
B_Q05aRU			8	ISCED 6	000000100
B_Q05aRU			96	Valid skip	000000010
B_Q05aSE	15	Education - Formal qualification - Level	-1	Missing	00000000000001
B_Q05aSE			1	Not stated ok 1-6	00000000000000
B_Q05aSE			2	Not stated ok 7-9	10000000000000
B_Q05aSE			3	Grundskolekompetens	01000000000000
B_Q05aSE			4	Gymnasie fackskola y	00100000000000
B_Q05aSE			5	Gymnasie fackskola y	00010000000000
B_Q05aSE			6	Gymnasie fackskola y	00001000000000
B_Q05aSE			7	Vuxenutbildning mots	00000100000000
B_Q05aSE			8	Vuxenutbildning mots	00000010000000
B_Q05aSE			9	Eftergymnasial utbil	00000001000000
B_Q05aSE			10	Eftergymnasial utbil	00000000100000
B_Q05aSE			11	Eftergymnasial utbil	00000000010000
B_Q05aSE			12	Eftergymnasial utbil	00000000001000
B_Q05aSE			13	Forskarutbildning	00000000000100
B_Q05aSE			96	Valid skip	00000000000010
B_Q05aSK	12	Education - Formal qualification - Level	-1	Missing	00000000001
B_Q05aSK			1	Primary school 1-4.	00000000000
B_Q05aSK			2	Primary school 5.-9.	10000000000
B_Q05aSK			3	Secondary technical	01000000000
B_Q05aSK			4	Secondary technical	00100000000
B_Q05aSK			5	Secondary schools wi	00010000000
B_Q05aSK			6	Upper secondary scho	00001000000
B_Q05aSK			7	Pre-tertiary school,	00000100000
B_Q05aSK			8	Bachelor degree, Gra	00000010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q05aSK			9	Master degree	00000001000
B_Q05aSK			10	PhD studies, Second	00000000100
B_Q05aSK			96	Valid skip	00000000010
B_Q05aUK1	27	Education - Formal qualification - Level	-1	Missing	00000000000000000000000001
B_Q05aUK1			1	Degree level qualifi	00000000000000000000000000
B_Q05aUK1			2	Diploma in higher ed	10000000000000000000000000
B_Q05aUK1			3	HNC/HND	01000000000000000000000000
B_Q05aUK1			4	ONC/OND	00100000000000000000000000
B_Q05aUK1			5	EdExcel/LQL	00010000000000000000000000
B_Q05aUK1			6	SCOTVEC (Scotland)	00001000000000000000000000
B_Q05aUK1			7	Teaching qualificati	00000100000000000000000000
B_Q05aUK1			8	Nursing or other med	00000010000000000000000000
B_Q05aUK1			9	other Higher Educati	00000001000000000000000000
B_Q05aUK1			10	A Level/Vocational A	00000000100000000000000000
B_Q05aUK1			11	NVQ/SVQ	00000000010000000000000000
B_Q05aUK1			12	AS Level/Vocational	00000000001000000000000000
B_Q05aUK1			13	Access to HE	00000000000100000000000000
B_Q05aUK1			14	Advanced Highers (Sc	00000000000010000000000000
B_Q05aUK1			15	Highers (Scotland)	00000000000001000000000000
B_Q05aUK1			16	Intermediate 2 NQs (	00000000000000100000000000
B_Q05aUK1			17	Intermediate 1 NQs (	00000000000000010000000000
B_Q05aUK1			18	Access Level (Scotla	00000000000000001000000000
B_Q05aUK1			19	National Qualificati	00000000000000000100000000
B_Q05aUK1			20	GCSE/Vocational GCSE	00000000000000000010000000
B_Q05aUK1			21	RSA/OCR	00000000000000000001000000
B_Q05aUK1			22	City and Guilds	00000000000000000000100000
B_Q05aUK1			23	Key skills/Basic ski	00000000000000000000010000
B_Q05aUK1			24	Entry level qualific	00000000000000000000001000
B_Q05aUK1			25	Any other profession	00000000000000000000000100
B_Q05aUK1			96	Valid skip	00000000000000000000000010
B_Q05aUK2	7	Education - Formal qualification - NVQ/SVQ Level	-1	Missing	000001
B_Q05aUK2			1	Level 1	000000
B_Q05aUK2			2	Level 2	100000
B_Q05aUK2			3	Level 3	010000
B_Q05aUK2			4	Level 4	001000
B_Q05aUK2			5	Level 5	000100
B_Q05aUK2			6	Valid skip	000010
B_Q05aUK3	6	Education - Formal qualification - BTEC/EdExcel/LQ	-1	Missing	00001
B_Q05aUK3			1	A higher Level (leve	00000
B_Q05aUK3			2	National Certificate	10000
B_Q05aUK3			3	First Diploma or gen	01000
B_Q05aUK3			4	First certificate or	00100
B_Q05aUK3			6	Valid skip	00010
B_Q05aUK4	7	Education - Formal qualification - SCOTVEC Level	-1	Missing	000001
B_Q05aUK4			1	A higher Level (leve	000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q05aUK4			2	Full national certif	100000
B_Q05aUK4			3	A first diploma or g	010000
B_Q05aUK4			4	A first certificate	001000
B_Q05aUK4			5	Modules towards a Na	000100
B_Q05aUK4			6	Valid skip	000010
B_Q05aUK5	7	Education - Formal qualification - National Qualif	-1	Missing	000001
B_Q05aUK5			1	Access Level	000000
B_Q05aUK5			2	Intermediate 1	100000
B_Q05aUK5			3	Intermediate 2	010000
B_Q05aUK5			4	Higher	001000
B_Q05aUK5			5	Advanced Higher	000100
B_Q05aUK5			6	Valid skip	000010
B_Q05aUK6	6	Education - Formal qualification - RSA/OCR Level	-1	Missing	00001
B_Q05aUK6			1	a higher diploma	00000
B_Q05aUK6			2	an advanced diploma	10000
B_Q05aUK6			3	a diploma	01000
B_Q05aUK6			4	or some other RSA (i	00100
B_Q05aUK6			6	Valid skip	00010
B_Q05aUK7	5	Education - Formal qualification - City & Guilds L	-1	Missing	0001
B_Q05aUK7			1	Advanced craft/part	0000
B_Q05aUK7			2	craft/part 2	1000
B_Q05aUK7			3	foundation/part 1	0100
B_Q05aUK7			6	Valid skip	0010
B_Q05aUK8	4	Education - Formal qualification - doing apprentic	-1	Missing	001
B_Q05aUK8			1	Yes	000
B_Q05aUK8			2	No	100
B_Q05aUK8			6	Valid skip	010
B_Q05aUS	12	Education - Formal qualification - Level	-1	Missing	00000000001
B_Q05aUS			1	Grades 1-6	00000000000
B_Q05aUS			2	Grades 7-9	10000000000
B_Q05aUS			3	High school diploma	01000000000
B_Q05aUS			4	Pre-associate educat	00100000000
B_Q05aUS			6	A certificate from a	00010000000
B_Q05aUS			7	Associate degree	00001000000
B_Q05aUS			8	Bachelor's degree (e	00000100000
B_Q05aUS			9	Master's degree (e.g	00000010000
B_Q05aUS			10	Professional degree	00000001000
B_Q05aUS			11	Doctorate degree (e.	00000000100
B_Q05aUS			96	Valid skip	00000000010
B_Q05bCZ	13	Education - Formal qualification - Area of study	-1	Missing	000000000001
B_Q05bCZ			1	General programmes	000000000000
B_Q05bCZ			2	Teacher training and	100000000000
B_Q05bCZ			3	Humanities, language	010000000000
B_Q05bCZ			4	Social sciences	001000000000
B_Q05bCZ			5	Business and law	000100000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q05bCZ			6	Science, mathematics	000010000000
B_Q05bCZ			7	Engineering, manufac	000001000000
B_Q05bCZ			8	Agriculture and vete	000000100000
B_Q05bCZ			9	Health	000000010000
B_Q05bCZ			10	Welfare	000000001000
B_Q05bCZ			11	Services	000000000100
B_Q05bCZ			96	Valid skip	000000000010
B_Q05bKO	12	KO_Education - Formal qualification - Area of stud	-1	Missing	00000000001
B_Q05bKO			1	General programmes	00000000000
B_Q05bKO			2	Teacher training and	10000000000
B_Q05bKO			3	Humanities, language	01000000000
B_Q05bKO			4	Social sciences, bus	00100000000
B_Q05bKO			5	Science, mathematics	00010000000
B_Q05bKO			6	Engineering, manufac	00001000000
B_Q05bKO			7	Agriculture and vete	00000100000
B_Q05bKO			8	Dental and medicine	00000010000
B_Q05bKO			9	Health and wellfare	00000001000
B_Q05bKO			10	Services	00000000100
B_Q05bKO			96	Valid skip	00000000010
B_Q05bNL	13	Education - Formal qualification - Area of study	-1	Missing	000000000001
B_Q05bNL			1	general programmes	000000000000
B_Q05bNL			2	teacher training, ed	100000000000
B_Q05bNL			3	humanities, language	010000000000
B_Q05bNL			4	social sciences, com	001000000000
B_Q05bNL			5	economy, business, m	000100000000
B_Q05bNL			6	law, civil service,	000010000000
B_Q05bNL			7	mathematics, natural	000001000000
B_Q05bNL			8	technics	000000100000
B_Q05bNL			9	agriculture, veterin	000000010000
B_Q05bNL			10	health, welfare, per	000000001000
B_Q05bNL			11	tourism, horeca, tra	000000000100
B_Q05bNL			96	Valid skip	000000000010
B_Q05bUK	21	Education - Formal qualification - Area of study	-1	Missing	00000000000000000001
B_Q05bUK			1	General programmes	00000000000000000000
B_Q05bUK			2	Medicine	10000000000000000000
B_Q05bUK			3	Medical related subj	01000000000000000000
B_Q05bUK			4	Biological Sciences	00100000000000000000
B_Q05bUK			5	Agricultural science	00010000000000000000
B_Q05bUK			6	Physical/Environment	00001000000000000000
B_Q05bUK			7	Mathematical Science	00000100000000000000
B_Q05bUK			8	Engineering	00000010000000000000
B_Q05bUK			9	Technology	00000001000000000000
B_Q05bUK			10	Architecture and rel	00000000100000000000
B_Q05bUK			11	Social Sciences (inc	00000000010000000000
B_Q05bUK			12	Business and Financi	00000000001000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q05bUK			13	Librarianship and In	00000000000100000000
B_Q05bUK			14	Linguistics, English	00000000000010000000
B_Q05bUK			15	European Languages	00000000000001000000
B_Q05bUK			16	Other languages	00000000000000100000
B_Q05bUK			17	Humanities	00000000000000010000
B_Q05bUK			18	Arts	00000000000000001000
B_Q05bUK			19	Education	00000000000000000100
B_Q05bUK			96	Valid skip	00000000000000000010
B_Q05cUSX1	5	Education - Formal qualification - Degree personal	-1	Missing	0001
B_Q05cUSX1			1	Yes, I studied as mu	0000
B_Q05cUSX1			2	Yes, but personal in	1000
B_Q05cUSX1			3	No	0100
B_Q05cUSX1			6	Valid skip	0010
B_Q05cUSX2	4	Education - Formal qualification - Degree personal	-1	Missing	001
B_Q05cUSX2			1	Personal interest	000
B_Q05cUSX2			2	Personal interest an	100
B_Q05cUSX2			6	Valid skip	010
B_Q10dEEX	6	Education - Formal qualification - Employed - Usef	-1	Missing	00001
B_Q10dEEX			1	Not useful at all	00000
B_Q10dEEX			2	Somewhat useful	10000
B_Q10dEEX			3	Moderately useful	01000
B_Q10dEEX			4	Very useful	00100
B_Q10dEEX			6	Valid skip	00010
B_Q11ATX1	4	Education - Kindergarten - NATIONAL	-1	Missing	001
B_Q11ATX1			1	Yes	000
B_Q11ATX1			2	No	100
B_Q11ATX1			6	Valid skip	010
B_Q11ATX3	6	Education - Lower secondary level - NATIONAL	-1	Missing	00001
B_Q11ATX3			1	Upper level of prima	00000
B_Q11ATX3			2	General secondary sc	10000
B_Q11ATX3			3	General secondary sc	01000
B_Q11ATX3			4	Other	00100
B_Q11ATX3			6	Valid skip	00010
B_Q11ATX4	6	Education - School leaving exam - NATIONAL	-1	Missing	00001
B_Q11ATX4			1	Not stated or inferr	00000
B_Q11ATX4			2	Vocational college	10000
B_Q11ATX4			3	No School leaving ex	01000
B_Q11ATX4			4	Other	00100
B_Q11ATX4			6	Valid skip	00010
B_Q11JPX1	6	Education - Formal qualification subsidy	-1	Missing	00001
B_Q11JPX1			1	Yes, totally	00000
B_Q11JPX1			2	Yes, partly	10000
B_Q11JPX1			3	No, not at all	01000
B_Q11JPX1			4	There were no such c	00100
B_Q11JPX1			6	Valid skip	00010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q11JPX2	11	Education - Formal qualification financial burden	-1	Missing	0000000001
B_Q11JPX2			1	Less than 50,000 yen	0000000000
B_Q11JPX2			2	50,000 to 99,999 yen	1000000000
B_Q11JPX2			3	100,000 to 199,999 y	0100000000
B_Q11JPX2			4	200,000 to 299,999 y	0010000000
B_Q11JPX2			5	300,000 to 499,999 y	0001000000
B_Q11JPX2			6	500,000 to 999,999 y	0000100000
B_Q11JPX2			7	1,000,000 to 1,499,9	0000010000
B_Q11JPX2			8	1,500,000 to 1,999,9	0000001000
B_Q11JPX2			9	2,000,000 yen or mor	0000000100
B_Q11JPX2			96	Valid skip	0000000010
B_Q11NLX	5	Education - Formal qualification - Initiative part	-1	Missing	0001
B_Q11NLX			1	respondent	0000
B_Q11NLX			2	employer	1000
B_Q11NLX			3	other	0100
B_Q11NLX			6	Valid skip	0010
B_Q13AU	6	Activities - Last year - Activity specified (AUS)	-1	Missing	00001
B_Q13AU			1	A correspondence or	00000
B_Q13AU			2	An organised session	10000
B_Q13AU			3	A seminar or worksho	01000
B_Q13AU			4	Other kind of course	00100
B_Q13AU			6	Valid skip	00010
B_Q14bUSX1	5	Activities - Last year - Activity Participation fo	-1	Missing	0001
B_Q14bUSX1			1	Yes, I participated	0000
B_Q14bUSX1			2	Yes, but personal in	1000
B_Q14bUSX1			3	No	0100
B_Q14bUSX1			6	Valid skip	0010
B_Q14bUSX2	4	Activities - Last year - Activity Participation ma	-1	Missing	001
B_Q14bUSX2			1	Personal interest	000
B_Q14bUSX2			2	Personal interest an	100
B_Q14bUSX2			6	Valid skip	010
B_Q16NLX	5	Activities - Last year - Initiative participation	-1	Missing	0001
B_Q16NLX			1	respondent	0000
B_Q16NLX			2	employer	1000
B_Q16NLX			3	other	0100
B_Q16NLX			6	Valid skip	0010
B_Q26aAU	4	Activities - Last year - Wanted but didn't start (	-1	Missing	001
B_Q26aAU			1	Yes	000
B_Q26aAU			2	No	100
B_Q26aAU			6	Valid skip	010
B_Q26bJPX	4	Activities - ICT skills	-1	Missing	001
B_Q26bJPX			1	Yes (Please specify.	000
B_Q26bJPX			2	No	100
B_Q26bJPX			6	Valid skip	010
B_Q26NLX1	4	Activities - Last year - Participation APL	-1	Missing	001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
B_Q26NLX1			1	Yes	000
B_Q26NLX1			2	No	100
B_Q26NLX1			6	Valid skip	010
B_Q27aUSX	4	Activites - Class - Class/tutor basic skills	-1	Missing	001
B_Q27aUSX			1	Yes	000
B_Q27aUSX			2	No	100
B_Q27aUSX			6	Valid skip	010
B_Q27bUSX	4	Activites - Class - Class/tutor GED	-1	Missing	001
B_Q27bUSX			1	Yes	000
B_Q27bUSX			2	No	100
B_Q27bUSX			6	Valid skip	010
B_Q27cUSX	4	Activites - Class - Class/tutor other equivalency	-1	Missing	001
B_Q27cUSX			1	Yes	000
B_Q27cUSX			2	No	100
B_Q27cUSX			6	Valid skip	010
B_Q27dUSX	5	Activites - Class - Class/tutor main reason	-1	Missing	0001
B_Q27dUSX			1	WORK-RELATED	0000
B_Q27dUSX			2	PERSONAL INTEREST	1000
B_Q27dUSX			3	BOTH EQUALLY	0100
B_Q27dUSX			6	Valid skip	0010
B_Q27eUSXb	8	Activites - Class - Class attendence, unit	-1	Missing	0000001
B_Q27eUSXb			1	Day	0000000
B_Q27eUSXb			2	Week	1000000
B_Q27eUSXb			3	Month	0100000
B_Q27eUSXb			4	Semester	0010000
B_Q27eUSXb			5	Quarter	0001000
B_Q27eUSXb			6	Other specify	0000100
B_Q27eUSXb			96	Valid skip	0000010
B_Q29aUSX	4	Activites - Apprentice - Was apprentice	-1	Missing	001
B_Q29aUSX			1	Yes	000
B_Q29aUSX			2	No	100
B_Q29aUSX			6	Valid skip	010
B_S26bEEX	7	Activities - Last year - Wanted but didn't start -	-1	Missing	000001
B_S26bEEX			1	I did not have infor	000000
B_S26bEEX			2	Temporary or chronic	100000
B_S26bEEX			3	The course was full	010000
B_S26bEEX			4	The expected benefit	001000
B_S26bEEX			5	Other	000100
B_S26bEEX			96	Valid skip	000010
B_S26bSEX	4	Activities - Last year - Wanted but didn't start -	-1	Missing	001
B_S26bSEX			1	Not stated or inferr	000
B_S26bSEX			2	Jag hade inte tid pg	100
B_S26bSEX			6	Valid skip	010
C_Q02aAU	5	Current status/work history - Last month - Looking	-1	Missing	0001
C_Q02aAU			1	Yes, full time work	0000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
C_Q02aAU			2	Yes, part-time work	1000
C_Q02aAU			3	No	0100
C_Q02aAU			6	Valid skip	0010
C_Q02aFRX	5	Current status/work history - Last month - Job see	-1	Missing	0001
C_Q02aFRX			1	Registred and having	0000
C_Q02aFRX			2	Registred but withou	1000
C_Q02aFRX			3	No registered	0100
C_Q02aFRX			6	Valid skip	0010
C_Q02aJPX	4	Current status/work history - Last month - Looking	-1	Missing	001
C_Q02aJPX			1	Yes	000
C_Q02aJPX			2	No	100
C_Q02aJPX			6	Valid skip	010
C_Q02aUK2	4	Current status/work history - Last month - Looking	-1	Missing	001
C_Q02aUK2			1	Yes	000
C_Q02aUK2			2	No	100
C_Q02aUK2			6	Valid skip	010
C_Q03DE_01	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03DE_01			1	Marked	000
C_Q03DE_01			2	Not marked	100
C_Q03DE_01			6	Valid skip	010
C_Q03DE_02	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03DE_02			1	Marked	000
C_Q03DE_02			2	Not marked	100
C_Q03DE_02			6	Valid skip	010
C_Q03DE_03	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03DE_03			1	Marked	000
C_Q03DE_03			2	Not marked	100
C_Q03DE_03			6	Valid skip	010
C_Q03DE_04	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03DE_04			1	Marked	000
C_Q03DE_04			2	Not marked	100
C_Q03DE_04			6	Valid skip	010
C_Q03DE_05	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03DE_05			1	Marked	000
C_Q03DE_05			2	Not marked	100
C_Q03DE_05			6	Valid skip	010
C_Q03DE_06	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03DE_06			1	Marked	000
C_Q03DE_06			2	Not marked	100
C_Q03DE_06			6	Valid skip	010
C_Q03DE_07	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03DE_07			1	Marked	000
C_Q03DE_07			2	Not marked	100
C_Q03DE_07			6	Valid skip	010
C_Q03DE_08	4	Current status/work history - Last month - Reason	-1	Missing	001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
C_Q03DE_08			1	Marked	000
C_Q03DE_08			2	Not marked	100
C_Q03DE_08			6	Valid skip	010
C_Q03DE_09	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03DE_09			1	Marked	000
C_Q03DE_09			2	Not marked	100
C_Q03DE_09			6	Valid skip	010
C_Q03DE_10	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03DE_10			1	Marked	000
C_Q03DE_10			2	Not marked	100
C_Q03DE_10			6	Valid skip	010
C_Q03DE_11	4	Current status/work history - Last month - Reason	-1	Missing	001
C_Q03DE_11			1	Marked	000
C_Q03DE_11			2	Not marked	100
C_Q03DE_11			6	Valid skip	010
C_Q04aAU	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04aAU			1	Yes	000
C_Q04aAU			2	No	100
C_Q04aAU			6	Valid skip	010
C_Q04aBE1	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04aBE1			1	Yes	000
C_Q04aBE1			2	No	100
C_Q04aBE1			6	Valid skip	010
C_Q04aBE2	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04aBE2			1	Yes	000
C_Q04aBE2			2	No	100
C_Q04aBE2			6	Valid skip	010
C_Q04aSEX1	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04aSEX1			1	Yes	000
C_Q04aSEX1			2	No	100
C_Q04aSEX1			6	Valid skip	010
C_Q04aSEX2	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04aSEX2			1	Yes	000
C_Q04aSEX2			2	No	100
C_Q04aSEX2			6	Valid skip	010
C_Q04aUK1	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04aUK1			1	Yes	000
C_Q04aUK1			2	No	100
C_Q04aUK1			6	Valid skip	010
C_Q04aUK2	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04aUK2			1	Yes	000
C_Q04aUK2			2	No	100
C_Q04aUK2			6	Valid skip	010
C_Q04aUK3	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04aUK3			1	Yes	000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
C_Q04aUK3			2	No	100
C_Q04aUK3			6	Valid skip	010
C_Q04bAU	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04bAU			1	Yes	000
C_Q04bAU			2	No	100
C_Q04bAU			6	Valid skip	010
C_Q04cAU	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04cAU			1	Yes	000
C_Q04cAU			2	No	100
C_Q04cAU			6	Valid skip	010
C_Q04dSE1	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04dSE1			1	Yes	000
C_Q04dSE1			2	No	100
C_Q04dSE1			6	Valid skip	010
C_Q04dSE2	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04dSE2			1	Yes	000
C_Q04dSE2			2	No	100
C_Q04dSE2			6	Valid skip	010
C_Q04eAU	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04eAU			1	Yes	000
C_Q04eAU			2	No	100
C_Q04eAU			6	Valid skip	010
C_Q04eUK2	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04eUK2			1	Yes	000
C_Q04eUK2			2	No	100
C_Q04eUK2			6	Valid skip	010
C_Q04fAU	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04fAU			1	Yes	000
C_Q04fAU			2	No	100
C_Q04fAU			6	Valid skip	010
C_Q04gAU	7	Current status/work history - Last month - Ways of	-1	Missing	000001
C_Q04gAU			1	Advertised or tender	000000
C_Q04gAU			2	Contacted friends/re	100000
C_Q04gAU			3	Other	010000
C_Q04gAU			4	Only looked in newsp	001000
C_Q04gAU			5	None of these	000100
C_Q04gAU			6	Valid skip	000010
C_Q04gIT1	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04gIT1			1	Yes	000
C_Q04gIT1			2	No	100
C_Q04gIT1			6	Valid skip	010
C_Q04gIT2	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04gIT2			1	Yes	000
C_Q04gIT2			2	No	100
C_Q04gIT2			6	Valid skip	010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
C_Q04iBEX1	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04iBEX1			1	Yes	000
C_Q04iBEX1			2	No	100
C_Q04iBEX1			6	Valid skip	010
C_Q04iBEX2	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04iBEX2			1	Yes	000
C_Q04iBEX2			2	No	100
C_Q04iBEX2			6	Valid skip	010
C_Q04iEE	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04iEE			1	Yes	000
C_Q04iEE			2	No	100
C_Q04iEE			6	Valid skip	010
C_Q04iJPX	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04iJPX			1	Yes	000
C_Q04iJPX			2	No	100
C_Q04iJPX			6	Valid skip	010
C_Q04iSEX	4	Current status/work history - Last month - Ways of	-1	Missing	001
C_Q04iSEX			1	Yes	000
C_Q04iSEX			2	No	100
C_Q04iSEX			6	Valid skip	010
C_Q05AU1X	5	Current status/work history - Not looking for work	-1	Missing	0001
C_Q05AU1X			1	Yes	0000
C_Q05AU1X			2	Maybe/It depends	1000
C_Q05AU1X			3	No	0100
C_Q05AU1X			6	Valid skip	0010
C_Q05AU3X	5	Current status/work history - If suitable childcar	-1	Missing	0001
C_Q05AU3X			1	Yes	0000
C_Q05AU3X			2	Maybe/It depends	1000
C_Q05AU3X			3	No	0100
C_Q05AU3X			6	Valid skip	0010
C_Q05AU4X	4	Current status/work history - Childcare available,	-1	Missing	001
C_Q05AU4X			1	Yes	000
C_Q05AU4X			2	No	100
C_Q05AU4X			6	Valid skip	010
C_Q05AUX	4	Current status/work history - Ability to start job	-1	Missing	001
C_Q05AUX			1	Yes	000
C_Q05AUX			2	No	100
C_Q05AUX			6	Valid skip	010
C_Q07CZ	11	Current status/work history - Subjective status	-1	Missing	0000000001
C_Q07CZ			1	Full-time employed (	0000000000
C_Q07CZ			2	Part-time employed (	1000000000
C_Q07CZ			3	Unemployed	0100000000
C_Q07CZ			4	Pupil, student	0010000000
C_Q07CZ			5	Apprentice, internsh	0001000000
C_Q07CZ			6	In retirement or ear	0000100000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
C_Q07CZ			7	Permanently disabled	0000010000
C_Q07CZ			8	Fulfilling domestic	0000001000
C_Q07CZ			9	Other	0000000100
C_Q07CZ			96	Valid skip	0000000010
C_Q07IE	11	Current status/work history - Subjective status	-1	Missing	0000000001
C_Q07IE			1	Full-time employed (	0000000000
C_Q07IE			2	Part-time employed (	1000000000
C_Q07IE			3	Unemployed	0100000000
C_Q07IE			4	Pupil, student	0010000000
C_Q07IE			5	Apprentice, internsh	0001000000
C_Q07IE			6	In retirement or ear	0000100000
C_Q07IE			7	Permanently disabled	0000010000
C_Q07IE			8	Fulfilling domestic	0000001000
C_Q07IE			9	Other	0000000100
C_Q07IE			96	Valid skip	0000000010
C_Q07JP	11	Current status/work history - Subjective status	-1	Missing	0000000001
C_Q07JP			1	Full-time employed (	0000000000
C_Q07JP			2	Part-time employed (	1000000000
C_Q07JP			3	Unemployed	0100000000
C_Q07JP			4	Pupil, student	0010000000
C_Q07JP			5	Apprentice, internsh	0001000000
C_Q07JP			6	In retirement or ear	0000100000
C_Q07JP			7	Permanently disabled	0000010000
C_Q07JP			8	Fulfilling domestic	0000001000
C_Q07JP			9	Other	0000000100
C_Q07JP			96	Valid skip	0000000010
C_Q07NL	11	Current status/work history - Subjective status	-1	Missing	0000000001
C_Q07NL			1	Full-time employed (	0000000000
C_Q07NL			2	Part-time employed (	1000000000
C_Q07NL			3	Unemployed	0100000000
C_Q07NL			4	Pupil, student	0010000000
C_Q07NL			5	Apprentice, internsh	0001000000
C_Q07NL			6	In retirement or ear	0000100000
C_Q07NL			7	Permanently disabled	0000010000
C_Q07NL			8	Fulfilling domestic	0000001000
C_Q07NL			9	Other	0000000100
C_Q07NL			96	Valid skip	0000000010
C_Q07NLX	4	Current status/work history - Combination working	-1	Missing	001
C_Q07NLX			1	Yes	000
C_Q07NLX			2	No	100
C_Q07NLX			6	Valid skip	010
C_Q08bca2	4	Current status/work history - Ever worked at a job	-1	Missing	001
C_Q08bca2			1	Yes	000
C_Q08bca2			2	No	100
C_Q08bca2			6	Valid skip	010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
C_Q10aAU	9	Current status/work history - Last 5 years - How m	-1	Missing	00000001
C_Q10aAU			1	1	00000000
C_Q10aAU			2	2	10000000
C_Q10aAU			3	3	01000000
C_Q10aAU			4	4	00100000
C_Q10aAU			5	5	00010000
C_Q10aAU			6	6	00001000
C_Q10aAU			7	7 or more	00000100
C_Q10aAU			96	Valid skip	00000010
C_Q10bCZ	4	Unemployment > 3 months	-1	Missing	001
C_Q10bCZ			1	Yes	000
C_Q10bCZ			2	No	100
C_Q10bCZ			6	Valid skip	010
C_Q11dkx1	5	new tasks in a job?	-1	Missing	0001
C_Q11dkx1			1	Yes	0000
C_Q11dkx1			2	No	1000
C_Q11dkx1			3	Never participated i	0100
C_Q11dkx1			96	Valid skip	0010
C_Q11dkx2	4	more responsibility in a job?	-1	Missing	001
C_Q11dkx2			1	Yes	000
C_Q11dkx2			2	No	100
C_Q11dkx2			96	Valid skip	010
C_Q11dkx3	4	higher income?	-1	Missing	001
C_Q11dkx3			1	Yes	000
C_Q11dkx3			2	No	100
C_Q11dkx3			96	Valid skip	010
C_Q11dkx4	4	better chances to stay in a job?	-1	Missing	001
C_Q11dkx4			1	Yes	000
C_Q11dkx4			2	No	100
C_Q11dkx4			96	Valid skip	010
C_Q11dkx5	4	better chances to get a new job?	-1	Missing	001
C_Q11dkx5			1	Yes	000
C_Q11dkx5			2	No	100
C_Q11dkx5			96	Valid skip	010
C_S07DEX	5	Participation in part time retirement scheme	-1	Missing	0001
C_S07DEX			1	Yes, still actively	0000
C_S07DEX			2	Yes, not actively wo	1000
C_S07DEX			3	No	0100
C_S07DEX			6	Valid skip	0010
Childunder13AU	5	Number of children under 13	-1	Missing	0001
Childunder13AU			1	One child under 13	0000
Childunder13AU			2	Two or more children	1000
Childunder13AU			3	No children under 13	0100
Childunder13AU			6	Valid skip	0010
CNT_BRTH_DEX	14	Country of birth - Respondent	-1	Missing	0000000000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
CNT_BRTH_DEX			276	Germany	0000000000000
CNT_BRTH_DEX			398	Kazakhstan	1000000000000
CNT_BRTH_DEX			616	Poland	0100000000000
CNT_BRTH_DEX			642	Romania	0010000000000
CNT_BRTH_DEX			643	Russian Federation	0001000000000
CNT_BRTH_DEX			792	Turkey	0000100000000
CNT_BRTH_DEX			1000	Europe	0000010000000
CNT_BRTH_DEX			2000	Africa	0000001000000
CNT_BRTH_DEX			3000	Americas	0000000100000
CNT_BRTH_DEX			4000	Asia	0000000010000
CNT_BRTH_DEX			5000	Australia and Oceani	0000000001000
CNT_BRTH_DEX			6000	Other	0000000000100
CNT_BRTH_DEX			9996	Valid skip	0000000000010
CNT_BRTHAU	5	Country of birth - Respondent (UN M49 numerical)	-1	Missing	0001
CNT_BRTHAU			36	Australia	0000
CNT_BRTHAU			826	Main English speakin	1000
CNT_BRTHAU			894	Other country	0100
CNT_BRTHAU			996	Valid skip	0010
CNT_HAU	4	Country in which highest qualification was gained	-1	Missing	001
CNT_HAU			36	Australia	000
CNT_HAU			894	Other country	100
CNT_HAU			996	Valid skip	010
D_D04	4	Current work - Employee or self-employed	-1	Missing	001
D_D04			1	Employee	000
D_D04			2	Self-employed	100
D_D04			6	Valid skip	010
D_D04AT	4	Current work - Employee or self-employed - NATIONA	-1	Missing	001
D_D04AT			1	Employee	000
D_D04AT			2	Self-employed	100
D_D04AT			6	Valid skip	010
D_Q01aFIX	4	Can I check, is your current job <INSERT JOB TITLE	-1	Missing	001
D_Q01aFIX			1	Yes	000
D_Q01aFIX			2	No	100
D_Q01aFIX			6	Valid skip	010
D_Q01aFR1	11	Current work - Job status	-1	Missing	0000000001
D_Q01aFR1			1	Civil servant workin	0000000000
D_Q01aFR1			2	Civil servant workin	1000000000
D_Q01aFR1			3	Employee on the Soci	0100000000
D_Q01aFR1			4	Employee of a public	0010000000
D_Q01aFR1			5	Employee of private	0001000000
D_Q01aFR1			6	Employee of and indi	0000100000
D_Q01aFR1			7	Employee in your own	0000010000
D_Q01aFR1			8	Running your own bnu	0000001000
D_Q01aFR1			9	Helping one of your	0000000100
D_Q01aFR1			96	Valid skip	0000000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q01aFR3	10	Current work - Job classification	-1	Missing	000000001
D_Q01aFR3			1	Unskilled industrial	000000000
D_Q01aFR3			2	Skilled industrial w	100000000
D_Q01aFR3			3	Technician	010000000
D_Q01aFR3			4	Civil servant with a	001000000
D_Q01aFR3			5	Civil servant with a	000100000
D_Q01aFR3			6	Civil servant with a	000010000
D_Q01aFR3			7	Civil servant with a	000001000
D_Q01aFR3			8	Other. Specify.	000000100
D_Q01aFR3			96	Valid skip	000000010
D_Q01aFR4	10	Current work - Job classification	-1	Missing	000000001
D_Q01aFR4			1	Unskilled industrial	000000000
D_Q01aFR4			2	Skilled industrial w	100000000
D_Q01aFR4			3	Technician	010000000
D_Q01aFR4			4	Foreman, salesman	001000000
D_Q01aFR4			5	Engineer, executive	000100000
D_Q01aFR4			6	Chief executive, top	000010000
D_Q01aFR4			7	Office clerck, sales	000001000
D_Q01aFR4			8	Other. Specify.	000000100
D_Q01aFR4			96	Valid skip	000000010
D_Q01aFR5	10	Current work - Job classification	-1	Missing	000000001
D_Q01aFR5			1	Director of your own	000000000
D_Q01aFR5			2	Leading manager of a	100000000
D_Q01aFR5			3	Free manager or rent	010000000
D_Q01aFR5			4	Minority manager	001000000
D_Q01aFR5			5	Associate	000100000
D_Q01aFR5			6	Partner in a busines	000010000
D_Q01aFR5			7	Other self-employed	000001000
D_Q01aFR5			8	Other. Specify.	000000100
D_Q01aFR5			96	Valid skip	000000010
D_Q01aFR6	12	Current work - Main task	-1	Missing	00000000001
D_Q01aFR6			1	Production, construc	00000000000
D_Q01aFR6			2	Repairing, maintaini	10000000000
D_Q01aFR6			3	Cleaning, caretaking	01000000000
D_Q01aFR6			4	Handing, logistics	00100000000
D_Q01aFR6			5	Secretary, reception	00010000000
D_Q01aFR6			6	Accounting, administ	00001000000
D_Q01aFR6			7	Sales and marketing	00000100000
D_Q01aFR6			8	Research and develop	00000010000
D_Q01aFR6			9	Education, healthcar	00000001000
D_Q01aFR6			10	Other. Specify.	00000000100
D_Q01aFR6			96	Valid skip	00000000010
D_Q01aNOX	4	Current work - Job title - Is registry correct	-1	Missing	001
D_Q01aNOX			1	Yes	000
D_Q01aNOX			2	No	100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q01aNOX			6	Valid skip	010
D_Q01aSEX	4	Current work - Job title - Register verification	-1	Missing	001
D_Q01aSEX			1	Yes	000
D_Q01aSEX			2	No	100
D_Q01aSEX			6	Valid skip	010
D_Q02aNOX	4	Current work - Kind of business, industry or servi	-1	Missing	001
D_Q02aNOX			1	Yes	000
D_Q02aNOX			2	No	100
D_Q02aNOX			6	Valid skip	010
D_Q02aSEX1	4	Current work - Verification	-1	Missing	001
D_Q02aSEX1			1	Yes	000
D_Q02aSEX1			2	No	100
D_Q02aSEX1			6	Valid skip	010
D_Q03US	5	Current work - Economic sector	-1	Missing	0001
D_Q03US			1	The private sector (	0000
D_Q03US			2	The public sector (f	1000
D_Q03US			3	A non-profit organis	0100
D_Q03US			6	Valid skip	0010
D_Q04AT1	8	Current work - Occupational status - NATIONAL	-1	Missing	0000001
D_Q04AT1			1	white-collar worker	0000000
D_Q04AT1			2	blue-collar worker	1000000
D_Q04AT1			3	magistrate	0100000
D_Q04AT1			4	Contract agent	0010000
D_Q04AT1			5	Freelancer	0001000
D_Q04AT1			6	self-employed	0000100
D_Q04AT1			96	Valid skip	0000010
D_Q04AT2	6	Current work - Degree of difficulty of the job - N	-1	Missing	00001
D_Q04AT2			1	easy tasks	00000
D_Q04AT2			2	average tasks	10000
D_Q04AT2			3	higher tasks	01000
D_Q04AT2			4	highly skilled tasks	00100
D_Q04AT2			6	Valid skip	00010
D_Q04AU	5	Current work - Work for Employer or in own busines	-1	Missing	0001
D_Q04AU			1	Employer	0000
D_Q04AU			2	Own business	1000
D_Q04AU			3	Other/Uncertain	0100
D_Q04AU			6	Valid skip	0010
D_Q04AU1	4	Current work - Form of payment - Wage or Salary	-1	Missing	001
D_Q04AU1			1	Wage/Salary	000
D_Q04AU1			2	Other/Uncertain	100
D_Q04AU1			6	Valid skip	010
D_Q04AU2	11	Current work - Payment or working arrangements	-1	Missing	0000000001
D_Q04AU2			1	Contractor/Subcontra	0000000000
D_Q04AU2			2	Own business/Partner	1000000000
D_Q04AU2			3	Commission only	0100000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q04AU2			4	Commission with reta	0010000000
D_Q04AU2			5	In a family business	0001000000
D_Q04AU2			6	Payment in kind	0000100000
D_Q04AU2			7	Paid by the price/it	0000010000
D_Q04AU2			8	Wage/salary earner	0000001000
D_Q04AU2			9	Other	0000000100
D_Q04AU2			96	Valid skip	0000000010
D_Q04AU3	4	Current work - Employees in business	-1	Missing	001
D_Q04AU3			1	Yes	000
D_Q04AU3			2	No	100
D_Q04AU3			6	Valid skip	010
D_Q04AU4	4	Current work - Is business incorporated	-1	Missing	001
D_Q04AU4			1	Yes	000
D_Q04AU4			2	No	100
D_Q04AU4			6	Valid skip	010
D_Q04EEX	8	Working - Planning own enterprise	-1	Missing	0000001
D_Q04EEX			1	I have not thought a	0000000
D_Q04EEX			2	I think about it	1000000
D_Q04EEX			3	I have gave up the i	0100000
D_Q04EEX			4	I am just about to s	0010000
D_Q04EEX			5	I was an entrepreneu	0001000
D_Q04EEX			6	I am actually an ent	0000100
D_Q04EEX			96	Valid skip	0000010
D_Q06aKO	8	KO_Current work - Amount of people working for emp	-1	Missing	0000001
D_Q06aKO			1	1 to 10 people	0000000
D_Q06aKO			2	11 to 50 people	1000000
D_Q06aKO			3	51 to 250 people	0100000
D_Q06aKO			4	251 to 300 people	0010000
D_Q06aKO			5	301 to 1000 people	0001000
D_Q06aKO			6	1001 people and over	0000100
D_Q06aKO			96	Valid skip	0000010
D_Q06cFRX	7	Current work - Size of compagny	-1	Missing	000001
D_Q06cFRX			1	1 to 10 people	000000
D_Q06cFRX			2	11 to 50 people	100000
D_Q06cFRX			3	51 to 250 people	010000
D_Q06cFRX			4	251 to 1000 people	001000
D_Q06cFRX			5	More than 1000 peopl	000100
D_Q06cFRX			6	Valid skip	000010
D_Q07bKO	8	KO_Current work - Employees working for you - Amou	-1	Missing	0000001
D_Q07bKO			1	1 to 10 people	0000000
D_Q07bKO			2	11 to 50 people	1000000
D_Q07bKO			3	51 to 250 people	0100000
D_Q07bKO			4	251 to 300 people	0010000
D_Q07bKO			5	301 to 1000 people	0001000
D_Q07bKO			6	1001 people and over	0000100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q07bKO			96	Valid skip	0000010
D_Q09ca1	10	Current work - Type of contract	-1	Missing	000000001
D_Q09ca1			1	A permanent contract	000000000
D_Q09ca1			2	A seasonal job	100000000
D_Q09ca1			3	A term or contract j	010000000
D_Q09ca1			4	A casual job	001000000
D_Q09ca1			5	Other temporary jobs	000100000
D_Q09ca1			6	An apprenticeship or	000010000
D_Q09ca1			7	No contract	000001000
D_Q09ca1			8	Other, please specif	000000100
D_Q09ca1			96	Valid skip	000000010
D_Q09CZ	7	Current work - Type of contract	-1	Missing	000001
D_Q09CZ			1	An indefinite contra	000000
D_Q09CZ			2	A fixed term contrac	100000
D_Q09CZ			3	A temporary employme	010000
D_Q09CZ			4	No contract	001000
D_Q09CZ			5	Other, please specif	000100
D_Q09CZ			6	Valid skip	000010
D_Q09DE	10	Current work - Type of contract	-1	Missing	000000001
D_Q09DE			1	An indefinite contra	000000000
D_Q09DE			2	A fixed term contrac	100000000
D_Q09DE			3	A temporary employme	010000000
D_Q09DE			4	An apprenticeship or	001000000
D_Q09DE			5	A honorary or freela	000100000
D_Q09DE			6	Seasonal contract	000010000
D_Q09DE			7	No written contract	000001000
D_Q09DE			8	Other	000000100
D_Q09DE			96	Valid skip	000000010
D_Q09EE	10	Current work - Type of contract	-1	Missing	000000001
D_Q09EE			1	Indefinite contract	000000000
D_Q09EE			2	Fixed term contract	100000000
D_Q09EE			3	A temporary subcontr	010000000
D_Q09EE			4	Indenture, incl publ	001000000
D_Q09EE			5	An apprenticeship co	000100000
D_Q09EE			6	A temporary contract	000010000
D_Q09EE			7	No contract	000001000
D_Q09EE			8	Other, please specif	000000100
D_Q09EE			96	Valid skip	000000010
D_Q09FR	9	Current work - Type of contract	-1	Missing	00000001
D_Q09FR			1	An indefinite contra	00000000
D_Q09FR			2	A fixed term contrac	10000000
D_Q09FR			3	A temporary employme	01000000
D_Q09FR			4	An apprenticeship	00100000
D_Q09FR			5	Training contract	00010000
D_Q09FR			6	No contract	00001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q09FR			7	Other. Specify.	00000100
D_Q09FR			96	Valid skip	00000010
D_Q09IT	9	Current work - Type of contract	-1	Missing	00000001
D_Q09IT			1	An indefinite contra	00000000
D_Q09IT			2	A fixed term contrac	10000000
D_Q09IT			3	A temporary employme	01000000
D_Q09IT			4	An apprenticeship or	00100000
D_Q09IT			5	Project-based contra	00010000
D_Q09IT			6	No contract	00001000
D_Q09IT			7	Other	00000100
D_Q09IT			96	Valid skip	00000010
D_Q09JP	12	Current work - Type of contract	-1	Missing	00000000001
D_Q09JP			1	Regular staff(indefi	00000000000
D_Q09JP			2	Regular staff(fixted	10000000000
D_Q09JP			3	A contract employee	01000000000
D_Q09JP			4	A part-time worker(i	00100000000
D_Q09JP			5	A part-time worker(f	00010000000
D_Q09JP			6	A temporary employme	00001000000
D_Q09JP			7	An entrusted employe	00000100000
D_Q09JP			8	An apprenticeship	00000010000
D_Q09JP			9	No contract	00000001000
D_Q09JP			10	Other	00000000100
D_Q09JP			96	Valid skip	00000000010
D_Q09KOX1	5	KO_Current work - Type of contract	-1	Missing	0001
D_Q09KOX1			1	A permanent worker	0000
D_Q09KOX1			2	A temporary worker	1000
D_Q09KOX1			3	A daily worker	0100
D_Q09KOX1			96	Valid skip	0010
D_Q09KOX2	4	KO_Current work - regular_irregural	-1	Missing	001
D_Q09KOX2			1	regular	000
D_Q09KOX2			2	irregular	100
D_Q09KOX2			96	Valid skip	010
D_Q09RU	7	Current work - Type of contract	-1	Missing	000001
D_Q09RU			1	An indefinite contra	000000
D_Q09RU			2	A fixed term contrac	100000
D_Q09RU			3	An apprenticeship or	010000
D_Q09RU			4	No contract	001000
D_Q09RU			5	Other	000100
D_Q09RU			96	Valid skip	000010
D_Q09SE	10	Current work - Type of contract	-1	Missing	000000001
D_Q09SE			1	Fast/tillsvidare	000000000
D_Q09SE			2	Fast/tillsvidare ll	100000000
D_Q09SE			3	Fast/tillsvidare ll	010000000
D_Q09SE			4	Fast/tillsvidare ll	001000000
D_Q09SE			5	Kallas vid behov	000100000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q09SE			6	Karling, praktik	000010000
D_Q09SE			7	Arbetsmarknadspoliti	000001000
D_Q09SE			8	Annan beskriv	000000100
D_Q09SE			96	Valid skip	000000010
D_Q10KOX3	6	KO_Current work - shift	-1	Missing	00001
D_Q10KOX3			1	No shift	00000
D_Q10KOX3			2	2 shifts	10000
D_Q10KOX3			3	3 shifts and over	01000
D_Q10KOX3			4	Work every other day	00100
D_Q10KOX3			6	Valid skip	00010
D_Q12aAT	18	Current work - Requirements - Education level - NA	-1	Missing	00000000000000001
D_Q12aAT			1	No compulsory school	00000000000000000
D_Q12aAT			2	Compulsory school	10000000000000000
D_Q12aAT			3	Apprenticeship	01000000000000000
D_Q12aAT			4	Vocational School (<	00100000000000000
D_Q12aAT			5	Vocational School (2	00010000000000000
D_Q12aAT			6	Nursing	00001000000000000
D_Q12aAT			7	Master craftsman's c	00000100000000000
D_Q12aAT			8	Academic Secondary S	00000010000000000
D_Q12aAT			9	Vocational college	00000001000000000
D_Q12aAT			10	Post-secondary cours	00000000100000000
D_Q12aAT			11	Post-secondary colle	00000000010000000
D_Q12aAT			12	University courses	00000000001000000
D_Q12aAT			13	University-Bachelor	00000000000100000
D_Q12aAT			14	University-Master	00000000000010000
D_Q12aAT			15	Post-graduate course	00000000000001000
D_Q12aAT			16	Doctoral Programme	00000000000000100
D_Q12aAT			96	Valid skip	00000000000000010
D_Q12aAU	17	Current work - Requirements - Education level	-1	Missing	0000000000000001
D_Q12aAU			1	Year 8 or below	0000000000000000
D_Q12aAU			2	Year 9 or equivalent	1000000000000000
D_Q12aAU			3	Year 10 or equivalen	0100000000000000
D_Q12aAU			4	Year 11 or equivalen	0010000000000000
D_Q12aAU			5	Year 12 or equivalen	0001000000000000
D_Q12aAU			6	Certificate I	0000100000000000
D_Q12aAU			7	Certificate II	0000010000000000
D_Q12aAU			8	Certificate III	0000001000000000
D_Q12aAU			9	Certificate IV	0000000100000000
D_Q12aAU			10	Diploma	0000000010000000
D_Q12aAU			11	Advanced Diploma and	0000000001000000
D_Q12aAU			12	Bachelor degree (inc	0000000000100000
D_Q12aAU			13	Graduate Diploma or	0000000000010000
D_Q12aAU			14	Masters	0000000000001000
D_Q12aAU			15	Doctorate	0000000000000100
D_Q12aAU			96	Valid skip	0000000000000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q12aBE	13	Current work - Requirements - Education level	-1	Missing	000000000001
D_Q12aBE			1	No formal qualificat	000000000000
D_Q12aBE			2	ISCED 1	100000000000
D_Q12aBE			3	ISCED 2	010000000000
D_Q12aBE			4	ISCED 3C 2 years or	001000000000
D_Q12aBE			5	ISCED 3A-B	000100000000
D_Q12aBE			6	ISCED 3 (without dis	000010000000
D_Q12aBE			7	ISCED 4A-B	000001000000
D_Q12aBE			8	ISCED 5B	000000100000
D_Q12aBE			9	ISCED 5A, bachelor d	000000010000
D_Q12aBE			10	ISCED 5A, master deg	000000001000
D_Q12aBE			11	ISCED 6	000000000100
D_Q12aBE			96	Valid skip	000000000010
D_Q12aca	15	Current work - Requirements - Education level	-1	Missing	00000000000001
D_Q12aca			1	No formal education	00000000000000
D_Q12aca			2	Grade 6	10000000000000
D_Q12aca			3	Less than high schoo	01000000000000
D_Q12aca			4	High school diploma	00100000000000
D_Q12aca			5	Trade/vocational cer	00010000000000
D_Q12aca			6	Apprenticeship certi	00001000000000
D_Q12aca			7	Non-university certi	00000100000000
D_Q12aca			8	University certifica	00000010000000
D_Q12aca			9	Bachelor's degree	00000001000000
D_Q12aca			10	University certifica	00000000100000
D_Q12aca			11	First professional d	00000000010000
D_Q12aca			12	Master's	00000000001000
D_Q12aca			13	Ph.D.	00000000000100
D_Q12aca			96	Valid skip	00000000000010
D_Q12aCY	10	Current work - Requirements - Education level	-1	Missing	000000001
D_Q12aCY			1	I never went to scho	000000000
D_Q12aCY			2	Primary school	100000000
D_Q12aCY			3	Public/Private Secon	010000000
D_Q12aCY			4	High School/Vocation	001000000
D_Q12aCY			5	Non-Univ. Degree/Dip	000100000
D_Q12aCY			6	Undergraduate degree	000010000
D_Q12aCY			7	Postgraduate degree,	000001000
D_Q12aCY			8	Doctorate	000000100
D_Q12aCY			96	Valid skip	000000010
D_Q12aCZ	15	Current work - Requirements - Education level	-1	Missing	00000000000001
D_Q12aCZ			1	No formal education	00000000000000
D_Q12aCZ			2	First level of basic	10000000000000
D_Q12aCZ			3	basic ISCED 2	01000000000000
D_Q12aCZ			4	vocational without m	00100000000000
D_Q12aCZ			5	vocational without m	00010000000000
D_Q12aCZ			6	ISCED 3A vocational	00001000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q12aCZ			7	ISCED 3A technical w	00000100000000
D_Q12aCZ			8	ISCED 3A general wit	00000010000000
D_Q12aCZ			9	ISCED 4 follow-up co	00000001000000
D_Q12aCZ			10	ISCED 5B higher prof	00000000100000
D_Q12aCZ			11	ISCED 5A, bachelor	00000000010000
D_Q12aCZ			12	ISCED 5A, master	00000000001000
D_Q12aCZ			13	ISCED 6, post gradua	00000000000100
D_Q12aCZ			96	Valid skip	00000000000010
D_Q12aDE1	13	Current work - Requirements - Professional qualifi	-1	Missing	000000000001
D_Q12aDE1			1	No professional qual	000000000000
D_Q12aDE1			2	Apprenticeship (Lehr	100000000000
D_Q12aDE1			3	Basic vocational tra	010000000000
D_Q12aDE1			4	Training at Fachschu	001000000000
D_Q12aDE1			5	Berufsakademie, Fach	000100000000
D_Q12aDE1			6	Bachelor at Fachhoch	000010000000
D_Q12aDE1			7	Master/Diplom at Fac	000001000000
D_Q12aDE1			8	Bachelor at universi	000000100000
D_Q12aDE1			9	Master/Diplom at uni	000000010000
D_Q12aDE1			10	Doctorate	000000001000
D_Q12aDE1			11	Another professional	000000000100
D_Q12aDE1			96	Valid skip	000000000010
D_Q12aDE2	7	Current work - Requirements - School qualification	-1	Missing	000001
D_Q12aDE2			1	Hauptschulabschluss	000000
D_Q12aDE2			2	Realschulabschluss (	100000
D_Q12aDE2			3	Fachhochschulreife,	010000
D_Q12aDE2			4	Abitur/EOS (General	001000
D_Q12aDE2			5	Another school leavi	000100
D_Q12aDE2			6	Valid skip	000010
D_Q12aDK	16	Current work - Requirements - Education level	-1	Missing	000000000000001
D_Q12aDK			1	No formal education	000000000000000
D_Q12aDK			2	Primary school, grad	100000000000000
D_Q12aDK			3	Lower secondary, gra	010000000000000
D_Q12aDK			4	Upper secondary voca	001000000000000
D_Q12aDK			5	Upper secondary voca	000100000000000
D_Q12aDK			6	Upper secondary gene	000010000000000
D_Q12aDK			7	Upper secondary unde	000001000000000
D_Q12aDK			8	Post secondary short	000000100000000
D_Q12aDK			9	Post secondary entra	000000010000000
D_Q12aDK			10	Post secondary non t	000000001000000
D_Q12aDK			11	Tertiary not researc	000000000100000
D_Q12aDK			12	Bachelor degree	000000000010000
D_Q12aDK			13	Master degree	000000000001000
D_Q12aDK			14	Ph.d or otther resea	000000000000100
D_Q12aDK			96	Valid skip	000000000000010
D_Q12aEE	20	Current work - Requirements - Education level	-1	Missing	0000000000000000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q12aEE			1	Without primary educ	0000000000000000000
D_Q12aEE			2	Primary education	1000000000000000000
D_Q12aEE			3	Basic education	0100000000000000000
D_Q12aEE			4	General secondary ed	0010000000000000000
D_Q12aEE			5	Vocational education	0001000000000000000
D_Q12aEE			6	Vocational education	0000100000000000000
D_Q12aEE			7	Vocational education	0000010000000000000
D_Q12aEE			8	Vocational secondary	0000001000000000000
D_Q12aEE			9	Secondary specialise	0000000100000000000
D_Q12aEE			10	Vocational secondary	0000000010000000000
D_Q12aEE			11	Secondary specialise	0000000001000000000
D_Q12aEE			12	Applied higher educa	0000000000100000000
D_Q12aEE			13	Bachelor's degree (3	0000000000010000000
D_Q12aEE			14	Bachelor's degree (4	0000000000001000000
D_Q12aEE			15	Higher education (st	0000000000000100000
D_Q12aEE			16	Master's degree (3+2	0000000000000010000
D_Q12aEE			17	Master's degree (4+2	0000000000000001000
D_Q12aEE			18	Doctoral degree (inc	0000000000000000100
D_Q12aEE			96	Valid skip	0000000000000000010
D_Q12aES	13	Current work - Requirements - Education level	-1	Missing	000000000001
D_Q12aES			1	Not stated or inferr	000000000000
D_Q12aES			2	Not stated or inferr	100000000000
D_Q12aES			3	Not stated or inferr	010000000000
D_Q12aES			4	Not stated or inferr	001000000000
D_Q12aES			5	Not stated or inferr	000100000000
D_Q12aES			6	Bachillerato,. y sim	000010000000
D_Q12aES			7	Pruebas de acceso a	000001000000
D_Q12aES			8	Pruebas de acceso a	000000100000
D_Q12aES			9	Pruebas de acceso a	000000010000
D_Q12aES			10	Pruebas de aster y e	000000001000
D_Q12aES			11	Programas de doctora	000000000100
D_Q12aES			96	Valid skip	000000000010
D_Q12aFI	13	Current work - Requirements - Education level	-1	Missing	000000000001
D_Q12aFI			1	No formal qualificat	000000000000
D_Q12aFI			2	ISCED 1	100000000000
D_Q12aFI			3	ISCED 2	010000000000
D_Q12aFI			4	Upper secondary voca	001000000000
D_Q12aFI			5	General upper second	000100000000
D_Q12aFI			6	Specialist vocationa	000010000000
D_Q12aFI			7	Vocational post-seco	000001000000
D_Q12aFI			8	Polytechnic degree (	000000100000
D_Q12aFI			9	Bachelor's degree (I	000000010000
D_Q12aFI			10	Master's degree (ISC	000000001000
D_Q12aFI			11	Licentiate's and doc	000000000100
D_Q12aFI			96	Valid skip	000000000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q12aFR	16	Current work - Requirements - Education level	-1	Missing	000000000000001
D_Q12aFR			1	No formal qualificat	000000000000000
D_Q12aFR			2	ISCED 1	100000000000000
D_Q12aFR			3	ISCED 2	010000000000000
D_Q12aFR			4	ISCED 3C shorter tha	001000000000000
D_Q12aFR			5	ISCED 3C 2 years or	000100000000000
D_Q12aFR			6	ISCED 3A-B	000010000000000
D_Q12aFR			7	ISCED 3 (without dis	000001000000000
D_Q12aFR			8	ISCED 4C	000000100000000
D_Q12aFR			9	ISCED 4A-B	000000010000000
D_Q12aFR			10	ISCED 4 (without dis	000000001000000
D_Q12aFR			11	ISCED 5B	000000000100000
D_Q12aFR			12	ISCED 5A, bachelor d	000000000010000
D_Q12aFR			13	ISCED 5A, master deg	000000000001000
D_Q12aFR			14	ISCED 6	000000000000100
D_Q12aFR			96	Valid skip	000000000000010
D_Q12aIE	15	Current work - Requirements - Education level	-1	Missing	00000000000001
D_Q12aIE			1	No formal education	00000000000000
D_Q12aIE			2	Primary education (o	10000000000000
D_Q12aIE			3	Secondary 1 (Junior/	01000000000000
D_Q12aIE			4	Transition year prog	00100000000000
D_Q12aIE			5	Secondary 2 (Leaving	00010000000000
D_Q12aIE			6	Technical or Vocatio	00001000000000
D_Q12aIE			7	Advanced Certificate	00000100000000
D_Q12aIE			8	Higher Certificate (	00000010000000
D_Q12aIE			9	Diploma (e.g. Nation	00000001000000
D_Q12aIE			10	Honours Bachelor Deg	00000000100000
D_Q12aIE			11	Professional (Honour	00000000010000
D_Q12aIE			12	Post-Graduate (e.g.	00000000001000
D_Q12aIE			13	Doctorate or higher	00000000000100
D_Q12aIE			96	Valid skip	00000000000010
D_Q12aIT	13	Current work - Requirements - Education level	-1	Missing	000000000001
D_Q12aIT			1	Non formal education	000000000000
D_Q12aIT			2	Primary education or	100000000000
D_Q12aIT			3	Lower secondary or s	010000000000
D_Q12aIT			4	Professional qualifi	001000000000
D_Q12aIT			5	Upper secondary educ	000100000000
D_Q12aIT			6	Post-secondary non t	000010000000
D_Q12aIT			7	Music Conservatory D	000001000000
D_Q12aIT			8	First stage of terti	000000100000
D_Q12aIT			9	First or second leve	000000010000
D_Q12aIT			10	Specialisation degre	000000001000
D_Q12aIT			11	Research Doctoral de	000000000100
D_Q12aIT			96	Valid skip	000000000010
D_Q12aJP	16	Current work - Requirements - Education level	-1	Missing	000000000000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q12aJP			1	No formal school edu	000000000000000
D_Q12aJP			2	Elementary school	100000000000000
D_Q12aJP			3	Lower secondary scho	010000000000000
D_Q12aJP			4	Short-term course of	001000000000000
D_Q12aJP			5	Specialized course o	000100000000000
D_Q12aJP			6	General/integrated c	000010000000000
D_Q12aJP			7	Passed upper seconda	000001000000000
D_Q12aJP			8	Advanced course of u	000000100000000
D_Q12aJP			9	Regular/advanced cou	000000010000000
D_Q12aJP			10	Undergraduate progr	000000001000000
D_Q12aJP			11	Master's programs/Do	000000000100000
D_Q12aJP			12	Completed all work o	000000000010000
D_Q12aJP			13	Doctoral programs of	000000000001000
D_Q12aJP			14	Specialized training	000000000000100
D_Q12aJP			96	Valid skip	000000000000010
D_Q12aKO	13	KO_Current work - Requirements - Education level	-1	Missing	000000000001
D_Q12aKO			1	no formal education	000000000000
D_Q12aKO			2	Elementary school	100000000000
D_Q12aKO			3	Middle school	010000000000
D_Q12aKO			4	High school(college	001000000000
D_Q12aKO			5	High school(vocation	000100000000
D_Q12aKO			6	2-3 year college	000010000000
D_Q12aKO			7	4 year college(speci	000001000000
D_Q12aKO			8	4 year college(gener	000000100000
D_Q12aKO			9	Master's degree(spec	000000010000
D_Q12aKO			10	Master's degree(gene	000000001000
D_Q12aKO			11	Doctoral degree	000000000100
D_Q12aKO			96	Valid skip	000000000010
D_Q12aNL	18	Current work - Requirements - Education level	-1	Missing	00000000000000001
D_Q12aNL			1	no formal qualificat	00000000000000000
D_Q12aNL			2	primary education (i	10000000000000000
D_Q12aNL			3	sec education, first	01000000000000000
D_Q12aNL			4	sec education, first	00100000000000000
D_Q12aNL			5	secondary education,	00010000000000000
D_Q12aNL			6	secondary education,	00001000000000000
D_Q12aNL			7	secondary education,	00000100000000000
D_Q12aNL			8	secondary education,	00000010000000000
D_Q12aNL			9	secondary education,	00000001000000000
D_Q12aNL			10	secondary education,	00000000100000000
D_Q12aNL			11	secondary education,	00000000010000000
D_Q12aNL			12	tertiary education,	00000000001000000
D_Q12aNL			13	tertiary education,	00000000000100000
D_Q12aNL			14	tertiary education,	00000000000010000
D_Q12aNL			15	tertiary education,	00000000000001000
D_Q12aNL			16	tertiary education,	00000000000000100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q12aNL			96	Valid skip	00000000000000010
D_Q12aNO	13	Current work - Requirements - Education level	-1	Missing	0000000000001
D_Q12aNO			1	No formal qualificat	0000000000000
D_Q12aNO			2	ISCED 1	1000000000000
D_Q12aNO			3	ISCED 2	0100000000000
D_Q12aNO			4	ISCED 3C shorter tha	0010000000000
D_Q12aNO			5	ISCED 3C 2 years or	0001000000000
D_Q12aNO			6	ISCED 3A-B	0000100000000
D_Q12aNO			7	ISCED 4C	0000010000000
D_Q12aNO			8	ISCED 4A-B	0000001000000
D_Q12aNO			9	ISCED 5B	0000000100000
D_Q12aNO			10	ISCED 5A, bachelor d	0000000010000
D_Q12aNO			11	ISCED 5A, Master deg	0000000001000
D_Q12aNO			12	ISCED 6	0000000000100
D_Q12aPL	12	Current work - Requirements - Education level	-1	Missing	000000000001
D_Q12aPL			1	No formal qualificat	000000000000
D_Q12aPL			2	ISCED 1	100000000000
D_Q12aPL			3	ISCED 2	010000000000
D_Q12aPL			4	ISCED 3C	001000000000
D_Q12aPL			5	ISCED 3B	000100000000
D_Q12aPL			6	ISCED 3A	000010000000
D_Q12aPL			7	ISCED 4	000001000000
D_Q12aPL			8	BA, ISCED 5A (I degr	000000100000
D_Q12aPL			9	MA, ISCED 5A (II deg	000000010000
D_Q12aPL			10	ISCED 6	000000001000
D_Q12aPL			96	Valid skip	000000000100
D_Q12aRU	11	Current work - Requirements - Education level	-1	Missing	00000000001
D_Q12aRU			1	No formal qualificat	00000000000
D_Q12aRU			2	ISCED 1	10000000000
D_Q12aRU			3	ISCED 2	01000000000
D_Q12aRU			4	ISCED 3 (without dis	00100000000
D_Q12aRU			5	ISCED 4 (without dis	00010000000
D_Q12aRU			6	ISCED 5B	00001000000
D_Q12aRU			7	ISCED 5A, bachelor d	00000100000
D_Q12aRU			8	ISCED 5A, master deg	00000010000
D_Q12aRU			9	ISCED 6	00000001000
D_Q12aRU			96	Valid skip	00000000100
D_Q12aSE	14	Current work - Requirements - Education level	-1	Missing	00000000000001
D_Q12aSE			1	Not stated or inferr	00000000000000
D_Q12aSE			2	Not stated or inr	10000000000000
D_Q12aSE			3	Not stated or inrr	01000000000000
D_Q12aSE			4	Gymnasie eller yrkes	00100000000000
D_Q12aSE			5	Gymnasie eller yrkes	00010000000000
D_Q12aSE			6	Gymnasie eller yrkes	00001000000000
D_Q12aSE			7	Gymnasie eller yrkes	00000100000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q12aSE			8	Eftergymnasial utbil	0000001000000
D_Q12aSE			9	Eftergymnasial utbil	0000000100000
D_Q12aSE			10	Eftergymnasial utbil	0000000010000
D_Q12aSE			11	Eftergymnasial utbil	0000000001000
D_Q12aSE			12	Forskarutbildning	0000000000100
D_Q12aSE			96	Valid skip	0000000000010
D_Q12aSK	14	Current work - Requirements - Education level	-1	Missing	0000000000001
D_Q12aSK			1	Pre school education	0000000000000
D_Q12aSK			2	Primary school 1-4.	1000000000000
D_Q12aSK			3	Primary school 5.-9.	0100000000000
D_Q12aSK			4	Secondary technical	0010000000000
D_Q12aSK			5	Secondary technical	0001000000000
D_Q12aSK			6	Gymnasium (5 years o	0000100000000
D_Q12aSK			7	Secondary schools wi	0000010000000
D_Q12aSK			8	Upper secondary scho	0000001000000
D_Q12aSK			9	Pre-tertiary school,	0000000100000
D_Q12aSK			10	Bachelor degree, Gra	0000000010000
D_Q12aSK			11	Master degree	0000000001000
D_Q12aSK			12	PhD studies, Second	0000000000100
D_Q12aSK			96	Valid skip	0000000000010
D_Q12aUK1	29	Current work - Requirements - Education level	-1	Missing	0000000000000000000000000001
D_Q12aUK1			1	Degree level qualifi	0000000000000000000000000000
D_Q12aUK1			2	Diploma in higher ed	1000000000000000000000000000
D_Q12aUK1			3	HNC/HND	0100000000000000000000000000
D_Q12aUK1			4	ONC/OND	0010000000000000000000000000
D_Q12aUK1			5	BTEC, BEC, TEC or Ed	0001000000000000000000000000
D_Q12aUK1			6	SCOTVEC, SCOTEC, SCO	0000100000000000000000000000
D_Q12aUK1			7	Teaching qualificati	0000010000000000000000000000
D_Q12aUK1			8	Nursing or other med	0000001000000000000000000000
D_Q12aUK1			9	other Higher Educati	0000000100000000000000000000
D_Q12aUK1			10	A Level/Vocational A	0000000010000000000000000000
D_Q12aUK1			11	Highers (Scotland)	0000000001000000000000000000
D_Q12aUK1			12	NVQ/SVQ	0000000000100000000000000000
D_Q12aUK1			13	GNVQ/GSVQ	0000000000010000000000000000
D_Q12aUK1			14	AS Level/Vocational	0000000000001000000000000000
D_Q12aUK1			15	Advanced highers or	0000000000000100000000000000
D_Q12aUK1			16	Access to HE	0000000000000010000000000000
D_Q12aUK1			17	O Level/GCSE/Vocatio	0000000000000001000000000000
D_Q12aUK1			18	Intermediate 1 or 2	0000000000000000100000000000
D_Q12aUK1			19	Standard Grade or O	0000000000000000010000000000
D_Q12aUK1			20	National Qualificati	0000000000000000001000000000
D_Q12aUK1			21	RSA/OCR	0000000000000000000100000000
D_Q12aUK1			22	City and Guilds	0000000000000000000010000000
D_Q12aUK1			23	YT Certificate/YTP	0000000000000000000001000000
D_Q12aUK1			24	Key skills/Basic ski	0000000000000000000000100000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q12aUK1			25	Entry level qualific	00000000000000000000000010000
D_Q12aUK1			26	Any other profession	00000000000000000000000001000
D_Q12aUK1			27	No qualifications re	00000000000000000000000000100
D_Q12aUK1			96	Valid skip	00000000000000000000000000010
D_Q12aUK2	7	Current work - Requirements - Level of NVQ/SVQ	-1	Missing	000001
D_Q12aUK2			1	Level 1	000000
D_Q12aUK2			2	Level 2	100000
D_Q12aUK2			3	Level 3	010000
D_Q12aUK2			4	Level 4	001000
D_Q12aUK2			5	Level 5	000100
D_Q12aUK2			6	Valid skip	000010
D_Q12aUK3	6	Current work - Requirements - Level of BTEC/BEC/TE	-1	Missing	00001
D_Q12aUK3			1	A higher Level (leve	00000
D_Q12aUK3			2	National Certificate	10000
D_Q12aUK3			3	First Diploma or gen	01000
D_Q12aUK3			4	First certificate or	00100
D_Q12aUK3			6	Valid skip	00010
D_Q12aUK4	7	Current work - Requirements - Level of SCOTVEC/SCO	-1	Missing	000001
D_Q12aUK4			1	A higher Level (leve	000000
D_Q12aUK4			2	Full national certif	100000
D_Q12aUK4			3	A first diploma or g	010000
D_Q12aUK4			4	A first certificate	001000
D_Q12aUK4			5	Modules towards a Na	000100
D_Q12aUK4			6	Valid skip	000010
D_Q12aUK5	7	Current work - Requirements - Level of GNVQ/GSVQ	-1	Missing	000001
D_Q12aUK5			1	Advanced level	000000
D_Q12aUK5			2	Full intermediate le	100000
D_Q12aUK5			3	Part 1 intermediate	010000
D_Q12aUK5			4	Full foundation leve	001000
D_Q12aUK5			5	Part 1 foundation le	000100
D_Q12aUK5			6	Valid skip	000010
D_Q12aUK6	7	Current work - Requirements - Level of National Qu	-1	Missing	000001
D_Q12aUK6			1	Access Level	000000
D_Q12aUK6			2	Intermediate 1	100000
D_Q12aUK6			3	Intermediate 2	010000
D_Q12aUK6			4	Higher	001000
D_Q12aUK6			5	Advanced Higher	000100
D_Q12aUK6			6	Valid skip	000010
D_Q12aUK7	6	Current work - Requirements - Level of RSA/OCR	-1	Missing	00001
D_Q12aUK7			1	a higher diploma	00000
D_Q12aUK7			2	an advanced diploma	10000
D_Q12aUK7			3	a diploma	01000
D_Q12aUK7			4	or some other RSA (i	00100
D_Q12aUK7			6	Valid skip	00010
D_Q12aUK8	5	Current work - Requirements - Level of City & Guil	-1	Missing	0001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q12aUK8			1	Advanced craft/part	0000
D_Q12aUK8			2	craft/part 2	1000
D_Q12aUK8			3	foundation/part 1	0100
D_Q12aUK8			6	Valid skip	0010
D_Q12aUKApp1	4	Current work - Requirements - Need apprenticeship	-1	Missing	001
D_Q12aUKApp1			1	Yes	000
D_Q12aUKApp1			2	No - it would requir	100
D_Q12aUKApp1			6	Valid skip	010
D_Q12aUS	13	Current work - Requirements - Education level	-1	Missing	000000000001
D_Q12aUS			1	Pre-primary or no sc	000000000000
D_Q12aUS			2	Grades 1-6	100000000000
D_Q12aUS			3	Grades 7-9	010000000000
D_Q12aUS			4	High school diploma	001000000000
D_Q12aUS			5	Pre-associate educat	000100000000
D_Q12aUS			7	A certificate from a	000010000000
D_Q12aUS			8	Associate degree	000001000000
D_Q12aUS			9	Bachelor's degree (e	000000100000
D_Q12aUS			10	Master's degree (e.g	000000010000
D_Q12aUS			11	Professional degree	000000001000
D_Q12aUS			12	Doctorate degree (e.	000000000100
D_Q12aUS			96	Valid skip	000000000010
D_Q13cATX1	7	Current work - Knowledge and skills - Utilized in	-1	Missing	000001
D_Q13cATX1			1	Not at all	000000
D_Q13cATX1			2	Very little	100000
D_Q13cATX1			3	To some extent	010000
D_Q13cATX1			4	To a high extent	001000
D_Q13cATX1			5	To a very high exten	000100
D_Q13cATX1			6	Valid skip	000010
D_Q13cATX2	7	Current work - Knowledge and skills - Learning Act	-1	Missing	000001
D_Q13cATX2			1	Not at all	000000
D_Q13cATX2			2	Very little	100000
D_Q13cATX2			3	To some extent	010000
D_Q13cATX2			4	To a high extent	001000
D_Q13cATX2			5	To a very high exten	000100
D_Q13cATX2			6	Valid skip	000010
D_Q13cATX3	7	Current work - Knowledge and skills - Learning Act	-1	Missing	000001
D_Q13cATX3			1	Not at all	000000
D_Q13cATX3			2	Very little	100000
D_Q13cATX3			3	To some extent	010000
D_Q13cATX3			4	To a high extent	001000
D_Q13cATX3			5	To a very high exten	000100
D_Q13cATX3			6	Valid skip	000010
D_Q13cATX4	7	Current work - Knowledge and skills - Learning Act	-1	Missing	000001
D_Q13cATX4			1	Not at all	000000
D_Q13cATX4			2	Very little	100000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q13cATX4			3	To some extent	010000
D_Q13cATX4			4	To a high extent	001000
D_Q13cATX4			5	To a very high exten	000100
D_Q13cATX4			6	Valid skip	000010
D_Q13cATX5	7	Current work - Knowledge and skills - Learning Act	-1	Missing	000001
D_Q13cATX5			1	Not at all	000000
D_Q13cATX5			2	Very little	100000
D_Q13cATX5			3	To some extent	010000
D_Q13cATX5			4	To a high extent	001000
D_Q13cATX5			5	To a very high exten	000100
D_Q13cATX5			6	Valid skip	000010
D_Q15aEEX	7	Current work - To what extent do you agree or disa	-1	Missing	000001
D_Q15aEEX			1	Strongly agree	000000
D_Q15aEEX			2	Agree	100000
D_Q15aEEX			3	Neither agree nor di	010000
D_Q15aEEX			4	Disagree	001000
D_Q15aEEX			5	Strongly disagree	000100
D_Q15aEEX			6	Valid skip	000010
D_Q15bEEX	7	Current work - To what extent do you agree or disa	-1	Missing	000001
D_Q15bEEX			1	Strongly agree	000000
D_Q15bEEX			2	Agree	100000
D_Q15bEEX			3	Neither agree nor di	010000
D_Q15bEEX			4	Disagree	001000
D_Q15bEEX			5	Strongly disagree	000100
D_Q15bEEX			6	Valid skip	000010
D_Q15cEEX	7	Current work - To what extent do you agree or disa	-1	Missing	000001
D_Q15cEEX			1	Strongly agree	000000
D_Q15cEEX			2	Agree	100000
D_Q15cEEX			3	Neither agree nor di	010000
D_Q15cEEX			4	Disagree	001000
D_Q15cEEX			5	Strongly disagree	000100
D_Q15cEEX			6	Valid skip	000010
D_Q15dEEX	7	Current work - To what extent do you agree or disa	-1	Missing	000001
D_Q15dEEX			1	Strongly agree	000000
D_Q15dEEX			2	Agree	100000
D_Q15dEEX			3	Neither agree nor di	010000
D_Q15dEEX			4	Disagree	001000
D_Q15dEEX			5	Strongly disagree	000100
D_Q15dEEX			6	Valid skip	000010
D_Q15eEEX	7	Current work - To what extent do you agree or disa	-1	Missing	000001
D_Q15eEEX			1	Strongly agree	000000
D_Q15eEEX			2	Agree	100000
D_Q15eEEX			3	Neither agree nor di	010000
D_Q15eEEX			4	Disagree	001000
D_Q15eEEX			5	Strongly disagree	000100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q15eEEX			6	Valid skip	000010
D_Q16aEEX	4	Current work - Earnings - Net or gross salary	-1	Missing	001
D_Q16aEEX			1	as sum that You get	000
D_Q16aEEX			2	together with taxes	100
D_Q16aEEX			6	Valid skip	010
D_Q16d1EE1	8	Current work - Earnings - Net pay per hour	-1	Missing	0000001
D_Q16d1EE1			1	up to 1,5 euro	0000000
D_Q16d1EE1			2	1,5-2 euro	1000000
D_Q16d1EE1			3	2,1-3 euro	0100000
D_Q16d1EE1			4	3,1-5 euro	0010000
D_Q16d1EE1			5	5,1-7 euro	0001000
D_Q16d1EE1			6	above 7 euro	0000100
D_Q16d1EE1			96	Valid skip	0000010
D_Q16d1EE2	8	Current work - Earnings - Gross pay per hour	-1	Missing	0000001
D_Q16d1EE2			1	up to 2 euro	0000000
D_Q16d1EE2			2	2,1-3 euro	1000000
D_Q16d1EE2			3	3,1-4 euro	0100000
D_Q16d1EE2			4	4,1-6 euro	0010000
D_Q16d1EE2			5	6,1-9 euro	0001000
D_Q16d1EE2			6	above 9 euro	0000100
D_Q16d1EE2			96	Valid skip	0000010
D_Q16d2EE1	8	Current work - Earnings - Net pay per day	-1	Missing	0000001
D_Q16d2EE1			1	up to 13 euro	0000000
D_Q16d2EE1			2	13-19 euro	1000000
D_Q16d2EE1			3	20-24 euro	0100000
D_Q16d2EE1			4	25-30 euro	0010000
D_Q16d2EE1			5	31-55 euro	0001000
D_Q16d2EE1			6	above 55 euro	0000100
D_Q16d2EE1			96	Valid skip	0000010
D_Q16d2EE2	8	Current work - Earnings - Gross pay per day	-1	Missing	0000001
D_Q16d2EE2			1	up to 15 euro	0000000
D_Q16d2EE2			2	15-20 euro	1000000
D_Q16d2EE2			3	21-30 euro	0100000
D_Q16d2EE2			4	31-45 euro	0010000
D_Q16d2EE2			5	46-70 euro	0001000
D_Q16d2EE2			6	above 70 euro	0000100
D_Q16d2EE2			96	Valid skip	0000010
D_Q16d3EE1	8	Current work - Earnings - Net pay per week	-1	Missing	0000001
D_Q16d3EE1			1	up to 70 euro	0000000
D_Q16d3EE1			2	70-90 euro	1000000
D_Q16d3EE1			3	91-130 euro	0100000
D_Q16d3EE1			4	131-190 euro	0010000
D_Q16d3EE1			5	191-260 euro	0001000
D_Q16d3EE1			6	above 260 euro	0000100
D_Q16d3EE1			96	Valid skip	0000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q16d3EE2	8	Current work - Earnings - Gross pay per week	-1	Missing	0000001
D_Q16d3EE2			1	up to 70 euro	0000000
D_Q16d3EE2			2	70-100 euro	1000000
D_Q16d3EE2			3	101-150 euro	0100000
D_Q16d3EE2			4	151-220 euro	0010000
D_Q16d3EE2			5	221-330 euro	0001000
D_Q16d3EE2			6	above 330 euro	0000100
D_Q16d3EE2			96	Valid skip	0000010
D_Q16d4EE1	8	Current work - Earnings - Net pay per 2 weeks	-1	Missing	0000001
D_Q16d4EE1			1	up to 130 euro	0000000
D_Q16d4EE1			2	130-180 euro	1000000
D_Q16d4EE1			3	181-260 euro	0100000
D_Q16d4EE1			4	261-370 euro	0010000
D_Q16d4EE1			5	371-540 euro	0001000
D_Q16d4EE1			6	above 540 euro	0000100
D_Q16d4EE1			96	Valid skip	0000010
D_Q16d4EE2	8	Current work - Earnings - Gross pay per 2 weeks	-1	Missing	0000001
D_Q16d4EE2			1	up to 140 euro	0000000
D_Q16d4EE2			2	140-200 euro	1000000
D_Q16d4EE2			3	201-300 euro	0100000
D_Q16d4EE2			4	301-450 euro	0010000
D_Q16d4EE2			5	451-650 euro	0001000
D_Q16d4EE2			6	above 650 euro	0000100
D_Q16d4EE2			96	Valid skip	0000010
D_Q16d5EE1	8	Current work - Earnings - Net pay per month	-1	Missing	0000001
D_Q16d5EE1			1	up to 270 euro	0000000
D_Q16d5EE1			2	270-400 euro	1000000
D_Q16d5EE1			3	401-550 euro	0100000
D_Q16d5EE1			4	551-800 euro	0010000
D_Q16d5EE1			5	801-1200 euro	0001000
D_Q16d5EE1			6	above 1200 euro	0000100
D_Q16d5EE1			96	Valid skip	0000010
D_Q16d5EE2	8	Current work - Earnings - Gross pay per month	-1	Missing	0000001
D_Q16d5EE2			1	up to 300 euro	0000000
D_Q16d5EE2			2	300-450 euro	1000000
D_Q16d5EE2			3	451-670 euro	0100000
D_Q16d5EE2			4	671-1000 euro	0010000
D_Q16d5EE2			5	1001-1450 euro	0001000
D_Q16d5EE2			6	above 1450 euro	0000100
D_Q16d5EE2			96	Valid skip	0000010
D_Q16d6EE1	8	Current work - Earnings - Net pay per year	-1	Missing	0000001
D_Q16d6EE1			1	up to 3300 euro	0000000
D_Q16d6EE1			2	3300-4600 euro	1000000
D_Q16d6EE1			3	4601-6600 euro	0100000
D_Q16d6EE1			4	6601-9600 euro	0010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q16d6EE1			5	9601-14000 euro	0001000
D_Q16d6EE1			6	above 14000 euro	0000100
D_Q16d6EE1			96	Valid skip	0000010
D_Q16d6EE2	8	Current work - Earnings - Gross pay per year	-1	Missing	0000001
D_Q16d6EE2			1	up to 3700 euro	0000000
D_Q16d6EE2			2	3700-5400 euro	1000000
D_Q16d6EE2			3	5401-8000 euro	0100000
D_Q16d6EE2			4	8001-12000 euro	0010000
D_Q16d6EE2			5	12001-17300 euro	0001000
D_Q16d6EE2			6	above 17300 euro	0000100
D_Q16d6EE2			96	Valid skip	0000010
D_Q16dFRX	4	Current work - Earnings - Broad categories - Gross	-1	Missing	001
D_Q16dFRX			1	Yes	000
D_Q16dFRX			2	No	100
D_Q16dFRX			6	Valid skip	010
D_Q16eATX	4	Current work - Earnings - Additional payments 13th	-1	Missing	001
D_Q16eATX			1	Yes	000
D_Q16eATX			2	No	100
D_Q16eATX			6	Valid skip	010
D_Q17aAT	4	Current work - Earnings - Additional payments - NA	-1	Missing	001
D_Q17aAT			1	Yes	000
D_Q17aAT			2	No	100
D_Q17aAT			6	Valid skip	010
D_Q17dEE1	5	Current work - Earnings - Additional payments - Br	-1	Missing	0001
D_Q17dEE1			1	less than 330 euro	0000
D_Q17dEE1			2	330-660 euro	1000
D_Q17dEE1			3	over 660 euro	0100
D_Q17dEE1			6	Valid skip	0010
D_Q17dEE2	5	Current work - Earnings - Additional payments - Br	-1	Missing	0001
D_Q17dEE2			1	less than 400 euro	0000
D_Q17dEE2			2	400-800 euro	1000
D_Q17dEE2			3	over 800 euro	0100
D_Q17dEE2			6	Valid skip	0010
D_Q18aAU1X	5	Current work - Earnings - Total business profit/lo	-1	Missing	0001
D_Q18aAU1X			1	Profit	0000
D_Q18aAU1X			2	Loss	1000
D_Q18aAU1X			3	Neither (nil income)	0100
D_Q18aAU1X			6	Valid skip	0010
D_Q18c1EE	8	Current work - Earnings - Broad categories - Total	-1	Missing	0000001
D_Q18c1EE			1	up to 300 euro	0000000
D_Q18c1EE			2	300-450 euro	1000000
D_Q18c1EE			3	451-670 euro	0100000
D_Q18c1EE			4	671-1000 euro	0010000
D_Q18c1EE			5	1001-1450 euro	0001000
D_Q18c1EE			6	above 1450 euro	0000100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
D_Q18c1EE			96	Valid skip	0000010
D_Q18c2EE	8	Current work - Earnings - Broad categories - Total	-1	Missing	0000001
D_Q18c2EE			1	up to 3700 euro	0000000
D_Q18c2EE			2	3700-5400 euro	1000000
D_Q18c2EE			3	5401-8000 euro	0100000
D_Q18c2EE			4	8001-12000 euro	0010000
D_Q18c2EE			5	12001-17300 euro	0001000
D_Q18c2EE			6	above 17300 euro	0000100
D_Q18c2EE			96	Valid skip	0000010
D_S16bAU	8	Current work - Earnings - Salary period	-1	Missing	0000001
D_S16bAU			1	Week	0000000
D_S16bAU			2	Fortnight	1000000
D_S16bAU			3	Four weeks	0100000
D_S16bAU			4	Calendar month	0010000
D_S16bAU			5	Year	0001000
D_S16bAU			6	Other (please specif	0000100
D_S16bAU			96	Valid skip	0000010
E_D04	4	Current work - Employee or self-employed	-1	Missing	001
E_D04			1	Employee	000
E_D04			2	Self-employed	100
E_D04			6	Valid skip	010
E_D04AT	4	Last job- Employee or self-employed - NATIONAL	-1	Missing	001
E_D04AT			1	Employee	000
E_D04AT			2	Self-employed	100
E_D04AT			6	Valid skip	010
E_Q01aFIX	4	Can I check, is your last job <INSERT JOB TITLE>?	-1	Missing	001
E_Q01aFIX			1	Yes	000
E_Q01aFIX			2	No	100
E_Q01aFIX			6	Valid skip	010
E_Q01aFR1	11	Last job - Job status	-1	Missing	0000000001
E_Q01aFR1			1	Civil servant workin	0000000000
E_Q01aFR1			2	Civil servant workin	1000000000
E_Q01aFR1			3	Employee on the Soci	0100000000
E_Q01aFR1			4	Employee of a public	0010000000
E_Q01aFR1			5	Employee of private	0001000000
E_Q01aFR1			6	Employee of and indi	0000100000
E_Q01aFR1			7	Employee in your own	0000010000
E_Q01aFR1			8	Running your own bnu	0000001000
E_Q01aFR1			9	Helping one of your	0000000100
E_Q01aFR1			96	Valid skip	0000000010
E_Q01aFR3	10	Last job - Job classification	-1	Missing	000000001
E_Q01aFR3			1	Unskilled industrial	000000000
E_Q01aFR3			2	Skilled industrial w	100000000
E_Q01aFR3			3	Technician	010000000
E_Q01aFR3			4	Civil servant with a	001000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
E_Q01aFR3			5	Civil servant with a	000100000
E_Q01aFR3			6	Civil servant with a	000010000
E_Q01aFR3			7	Civil servant with a	000001000
E_Q01aFR3			8	Other. Specify.	000000100
E_Q01aFR3			96	Valid skip	000000010
E_Q01aFR4	10	Last job - Job classification	-1	Missing	000000001
E_Q01aFR4			1	Unskilled industrial	000000000
E_Q01aFR4			2	Skilled industrial w	100000000
E_Q01aFR4			3	Technician	010000000
E_Q01aFR4			4	Foreman, salesman	001000000
E_Q01aFR4			5	Engineer, executive	000100000
E_Q01aFR4			6	Chief executive, top	000010000
E_Q01aFR4			7	Office clerck, sales	000001000
E_Q01aFR4			8	Other. Specify.	000000100
E_Q01aFR4			96	Valid skip	000000010
E_Q01aFR5	10	Last job - Job classification	-1	Missing	000000001
E_Q01aFR5			1	Director of your own	000000000
E_Q01aFR5			2	Leading manager of a	100000000
E_Q01aFR5			3	Free manager or rent	010000000
E_Q01aFR5			4	Minority manager	001000000
E_Q01aFR5			5	Associate	000100000
E_Q01aFR5			6	Partner in a busines	000010000
E_Q01aFR5			7	Other self-employed	000001000
E_Q01aFR5			8	Other. Specify.	000000100
E_Q01aFR5			96	Valid skip	000000010
E_Q01aFR6	12	Last job - Main task	-1	Missing	00000000001
E_Q01aFR6			1	Production, construc	00000000000
E_Q01aFR6			2	Repairing, maintaini	10000000000
E_Q01aFR6			3	Cleaning, caretaking	01000000000
E_Q01aFR6			4	Handing, logistics	00100000000
E_Q01aFR6			5	Secretary, reception	00010000000
E_Q01aFR6			6	Accounting, administ	00001000000
E_Q01aFR6			7	Sales and marketing	00000100000
E_Q01aFR6			8	Research and develop	00000010000
E_Q01aFR6			9	Education, healthcar	00000001000
E_Q01aFR6			10	Other. Specify.	00000000100
E_Q01aFR6			96	Valid skip	00000000010
E_Q01aNOX	4	Last job - Job title - Is registry correct	-1	Missing	001
E_Q01aNOX			1	Yes	000
E_Q01aNOX			2	No	100
E_Q01aNOX			6	Valid skip	010
E_Q02aNOX	4	Last job - Kind of business, industry or service	-1	Missing	001
E_Q02aNOX			1	Yes	000
E_Q02aNOX			2	No	100
E_Q02aNOX			6	Valid skip	010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
E_Q03US	5	Last job - Economic sector	-1	Missing	0001
E_Q03US			1	The private sector (	0000
E_Q03US			2	The public sector (f	1000
E_Q03US			3	A non-profit organis	0100
E_Q03US			6	Valid skip	0010
E_Q04AT1	8	Last job - Occupational status - NATIONAL	-1	Missing	0000001
E_Q04AT1			1	white-collar worker	0000000
E_Q04AT1			2	blue-collar worker	1000000
E_Q04AT1			3	magistrate	0100000
E_Q04AT1			4	Contract agent	0010000
E_Q04AT1			5	Freelancer	0001000
E_Q04AT1			6	self-employed	0000100
E_Q04AT1			96	Valid skip	0000010
E_Q04AT2	6	Last job - Degree of difficulty of the job - NATIO	-1	Missing	00001
E_Q04AT2			1	easy tasks	00000
E_Q04AT2			2	average tasks	10000
E_Q04AT2			3	higher tasks	01000
E_Q04AT2			4	highly skilled tasks	00100
E_Q04AT2			6	Valid skip	00010
E_Q04AU	5	Last job - Employee or self-employed	-1	Missing	0001
E_Q04AU			1	Employer	0000
E_Q04AU			2	Own business	1000
E_Q04AU			3	Other/Uncertain	0100
E_Q04AU			6	Valid skip	0010
E_Q04AU1	4	Last job - Form of payment - Wage or Salary	-1	Missing	001
E_Q04AU1			1	Wage/Salary	000
E_Q04AU1			2	Other/Uncertain	100
E_Q04AU1			6	Valid skip	010
E_Q04AU2	11	Last job - Payment or working arrangements	-1	Missing	0000000001
E_Q04AU2			1	Contractor/Subcontra	0000000000
E_Q04AU2			2	Own business/Partner	1000000000
E_Q04AU2			3	Commission only	0100000000
E_Q04AU2			4	Commission with reta	0010000000
E_Q04AU2			5	In a family business	0001000000
E_Q04AU2			6	Payment in kind	0000100000
E_Q04AU2			7	Paid by the price/it	0000010000
E_Q04AU2			8	Wage/salary earner	0000001000
E_Q04AU2			9	Other	0000000100
E_Q04AU2			96	Valid skip	0000000010
E_Q04AU3	4	Last job - Employees working for you	-1	Missing	001
E_Q04AU3			1	Yes	000
E_Q04AU3			2	No	100
E_Q04AU3			6	Valid skip	010
E_Q04AU4	4	Last job - Is business incorporated	-1	Missing	001
E_Q04AU4			1	Yes	000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
E_Q04AU4			2	No	100
E_Q04AU4			6	Valid skip	010
E_Q06bFRX	7	Last job - Size of compagny	-1	Missing	000001
E_Q06bFRX			1	1 to 10 people	000000
E_Q06bFRX			2	11 to 50 people	100000
E_Q06bFRX			3	51 to 250 people	010000
E_Q06bFRX			4	251 to 1000 people	001000
E_Q06bFRX			5	More than 1000 peopl	000100
E_Q06bFRX			6	Valid skip	000010
E_Q06KO	8	KO_Last job - Amount of people working for employe	-1	Missing	0000001
E_Q06KO			1	1 to 10 people	0000000
E_Q06KO			2	11 to 50 people	1000000
E_Q06KO			3	51 to 250 people	0100000
E_Q06KO			4	251 to 300 people	0010000
E_Q06KO			5	301 to 1000 people	0001000
E_Q06KO			6	1001 people and over	0000100
E_Q06KO			96	Valid skip	0000010
E_Q07bKO	8	KO_Last job - Employees working for you - Amount	-1	Missing	0000001
E_Q07bKO			1	1 to 10 people	0000000
E_Q07bKO			2	11 to 50 people	1000000
E_Q07bKO			3	51 to 250 people	0100000
E_Q07bKO			4	251 to 300 people	0010000
E_Q07bKO			5	301 to 1000 people	0001000
E_Q07bKO			6	1001 people and over	0000100
E_Q07bKO			96	Valid skip	0000010
E_Q08ca1	10	Last job - Type of contract	-1	Missing	000000001
E_Q08ca1			1	A permanent contract	000000000
E_Q08ca1			2	A seasonal job	100000000
E_Q08ca1			3	A term or contract j	010000000
E_Q08ca1			4	A casual job	001000000
E_Q08ca1			5	Other temporary jobs	000100000
E_Q08ca1			6	An apprenticeship or	000010000
E_Q08ca1			7	No contract	000001000
E_Q08ca1			8	Other, please specif	000000100
E_Q08ca1			96	Valid skip	000000010
E_Q08CZ	7	Last job - Type of contract	-1	Missing	000001
E_Q08CZ			1	An indefinite contra	000000
E_Q08CZ			2	A fixed term contrac	100000
E_Q08CZ			3	A temporary employme	010000
E_Q08CZ			4	No contract	001000
E_Q08CZ			5	Other, please specif	000100
E_Q08CZ			6	Valid skip	000010
E_Q08DE	10	Last job - Type of contract	-1	Missing	000000001
E_Q08DE			1	An indefinite contra	000000000
E_Q08DE			2	A fixed term contrac	100000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
E_Q08DE			3	A temporary employme	010000000
E_Q08DE			4	An apprenticeship or	001000000
E_Q08DE			5	A honorary or freela	000100000
E_Q08DE			6	Seasonal contract	000010000
E_Q08DE			7	No written contract	000001000
E_Q08DE			8	Other	000000100
E_Q08DE			96	Valid skip	000000010
E_Q08EE	10	Last job - Type of contract	-1	Missing	000000001
E_Q08EE			1	Indefinite contract	000000000
E_Q08EE			2	Fixed term contract	100000000
E_Q08EE			3	A temporary subcontr	010000000
E_Q08EE			4	Indenture, incl publ	001000000
E_Q08EE			5	An apprenticeship co	000100000
E_Q08EE			6	A temporary contract	000010000
E_Q08EE			7	No contract	000001000
E_Q08EE			8	Other, please specif	000000100
E_Q08EE			96	Valid skip	000000010
E_Q08FR	9	Last job - Type of contract	-1	Missing	00000001
E_Q08FR			1	An indefinite contra	00000000
E_Q08FR			2	A fixed term contrac	10000000
E_Q08FR			3	A temporary employme	01000000
E_Q08FR			4	An apprenticeship	00100000
E_Q08FR			5	Training contract	00010000
E_Q08FR			6	No contract	00001000
E_Q08FR			7	Other. Specify.	00000100
E_Q08FR			96	Valid skip	00000010
E_Q08IT	9	Last job - Type of contract	-1	Missing	00000001
E_Q08IT			1	An indefinite contra	00000000
E_Q08IT			2	A fixed term contrac	10000000
E_Q08IT			3	A temporary employme	01000000
E_Q08IT			4	An apprenticeship or	00100000
E_Q08IT			5	Project-based contra	00010000
E_Q08IT			6	No contract	00001000
E_Q08IT			7	Other	00000100
E_Q08IT			96	Valid skip	00000010
E_Q08JP	12	Last job- Type of contract	-1	Missing	00000000001
E_Q08JP			1	Regular staff(indefi	00000000000
E_Q08JP			2	Regular staff(fixted	10000000000
E_Q08JP			3	A contract employee	01000000000
E_Q08JP			4	A part-time worker(i	00100000000
E_Q08JP			5	A part-time worker(f	00010000000
E_Q08JP			6	A temporary employme	00001000000
E_Q08JP			7	An entrusted employe	00000100000
E_Q08JP			8	An apprenticeship	00000010000
E_Q08JP			9	No contract	00000001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
E_Q08JP			10	Other	00000000100
E_Q08JP			96	Valid skip	00000000010
E_Q08KOX1	5	KO_Last job - Type of contract	-1	Missing	0001
E_Q08KOX1			1	A permanent worker	0000
E_Q08KOX1			2	A temporary worker	1000
E_Q08KOX1			3	A daily worker	0100
E_Q08KOX1			96	Valid skip	0010
E_Q08KOX2	4	KO_Last job - regular_irregural	-1	Missing	001
E_Q08KOX2			1	regular	000
E_Q08KOX2			2	irregular	100
E_Q08KOX2			96	Valid skip	010
E_Q08RU	7	Last job - Type of contract	-1	Missing	000001
E_Q08RU			1	An indefinite contra	000000
E_Q08RU			2	A fixed term contrac	100000
E_Q08RU			3	An apprenticeship or	010000
E_Q08RU			4	No contract	001000
E_Q08RU			5	Other	000100
E_Q08RU			96	Valid skip	000010
E_Q08SE	10	Last job - Type of contract	-1	Missing	000000001
E_Q08SE			1	Fast/tillsvidare	000000000
E_Q08SE			2	Fast/tillsvidare ll	100000000
E_Q08SE			3	Fast/tillsvidare ll	010000000
E_Q08SE			4	Fast/tillsvidare ll	001000000
E_Q08SE			5	Kallas vid behov	000100000
E_Q08SE			6	Karling, praktik	000010000
E_Q08SE			7	Arbetsmarknadspoliti	000001000
E_Q08SE			8	Annan beskriv	000000100
E_Q08SE			96	Valid skip	000000010
E_Q09KOX3	6	KO_Last job - shift	-1	Missing	00001
E_Q09KOX3			1	No shift	00000
E_Q09KOX3			2	2 shifts	10000
E_Q09KOX3			3	3 shifts and over	01000
E_Q09KOX3			4	Work every other day	00100
E_Q09KOX3			6	Valid skip	00010
E_Q10AT	13	Last job - Reason for end of job -NATIONAL	-1	Missing	000000000001
E_Q10AT			1	I was dismissed	000000000000
E_Q10AT			2	I was made redundant	100000000000
E_Q10AT			3	It was a temporary j	010000000000
E_Q10AT			4	I resigned	001000000000
E_Q10AT			5	I gave up work for h	000100000000
E_Q10AT			6	I took early retirem	000010000000
E_Q10AT			7	I retired (at or aft	000001000000
E_Q10AT			8	I gave up work becau	000000100000
E_Q10AT			9	I gave up work in or	000000010000
E_Q10AT			10	I went to military s	000000001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
E_Q10AT			11	I left for some othe	000000000100
E_Q10AT			96	Valid skip	000000000010
E_Q10JPX	8	Last job - Reason for end of job	-1	Missing	0000001
E_Q10JPX			1	Slumping business or	0000000
E_Q10JPX			2	Just temporary job	1000000
E_Q10JPX			3	Low income	0100000
E_Q10JPX			4	Bad working conditio	0010000
E_Q10JPX			5	I am not suited for	0001000
E_Q10JPX			6	Other reason	0000100
E_Q10JPX			96	Valid skip	0000010
F_Q01aca1_01	4	Skill use work - Language used most often at work	-1	Missing	001
F_Q01aca1_01			1	Marked	000
F_Q01aca1_01			2	Not marked	100
F_Q01aca1_01			6	Valid skip	010
F_Q01aca1_02	4	Skill use work - Language used most often at work	-1	Missing	001
F_Q01aca1_02			1	Marked	000
F_Q01aca1_02			2	Not marked	100
F_Q01aca1_02			6	Valid skip	010
F_Q01aca1_03	4	Skill use work - Language used most often at work	-1	Missing	001
F_Q01aca1_03			1	Marked	000
F_Q01aca1_03			2	Not marked	100
F_Q01aca1_03			6	Valid skip	010
F_Q07bEEX1	7	Skill use to establish an enterprise - Have experi	-1	Missing	000001
F_Q07bEEX1			1	Not at all	000000
F_Q07bEEX1			2	Very little	100000
F_Q07bEEX1			3	To some extent	010000
F_Q07bEEX1			4	To a high extent	001000
F_Q07bEEX1			5	To a very high exten	000100
F_Q07bEEX1			6	Valid skip	000010
F_Q07bEEX2	7	Skill use to establish an enterprise - Business pl	-1	Missing	000001
F_Q07bEEX2			1	Not at all	000000
F_Q07bEEX2			2	Very little	100000
F_Q07bEEX2			3	To some extent	010000
F_Q07bEEX2			4	To a high extent	001000
F_Q07bEEX2			5	To a very high exten	000100
F_Q07bEEX2			6	Valid skip	000010
F_Q07bEEX4	7	Skill use to establish an enterprise - Know whom t	-1	Missing	000001
F_Q07bEEX4			1	Not at all	000000
F_Q07bEEX4			2	Very little	100000
F_Q07bEEX4			3	To some extent	010000
F_Q07bEEX4			4	To a high extent	001000
F_Q07bEEX4			5	To a very high exten	000100
F_Q07bEEX4			6	Valid skip	000010
F_Q07bEEX9	7	Skill use to establish an enterprise - Have experi	-1	Missing	000001
F_Q07bEEX9			1	Not at all	000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
F_Q07bEEX9			2	Very little	100000
F_Q07bEEX9			3	To some extent	010000
F_Q07bEEX9			4	To a high extent	001000
F_Q07bEEX9			5	To a very high exten	000100
F_Q07bEEX9			6	Valid skip	000010
G_Q01hKOX	7	KO_Skill use work - Literacy - Read diagrams maps	-1	Missing	000001
G_Q01hKOX			1	Less than 1 page	000000
G_Q01hKOX			2	2-5 pages	100000
G_Q01hKOX			3	6-10 pages	010000
G_Q01hKOX			4	11-25 pages	001000
G_Q01hKOX			5	26 pages and over	000100
G_Q01hKOX			6	Valid skip	000010
G_Q02dKOX	7	KO_Skill use work - Literacy - Fill in forms	-1	Missing	000001
G_Q02dKOX			1	Less than 1 page	000000
G_Q02dKOX			2	2-5 pages	100000
G_Q02dKOX			3	6-10 pages	010000
G_Q02dKOX			4	11-25 pages	001000
G_Q02dKOX			5	26 pages and over	000100
G_Q02dKOX			6	Valid skip	000010
G_Q04USX	4	Skill Use Work - ICT - Computer last job	-1	Missing	001
G_Q04USX			1	Yes	000
G_Q04USX			2	No	100
G_Q04USX			6	Valid skip	010
H_Q01cca4	7	Skill use everyday life - Literacy - Read newspape	-1	Missing	000001
H_Q01cca4			1	Never	000000
H_Q01cca4			2	Less than once a mon	100000
H_Q01cca4			3	Less than once a wee	010000
H_Q01cca4			4	At least once a week	001000
H_Q01cca4			5	Every day	000100
H_Q01cca4			6	Valid skip	000010
H_Q01eca4	7	Skill use everyday life - Literacy - Read books in	-1	Missing	000001
H_Q01eca4			1	Never	000000
H_Q01eca4			2	Less than once a mon	100000
H_Q01eca4			3	Less than once a wee	010000
H_Q01eca4			4	At least once a week	001000
H_Q01eca4			5	Every day	000100
H_Q01eca4			6	Valid skip	000010
I_Q010bUSX1	4	About yourself - Health - Have medical insurance	-1	Missing	001
I_Q010bUSX1			1	Yes	000
I_Q010bUSX1			2	No	100
I_Q010bUSX1			6	Valid skip	010
I_Q05aEEX	7	About yourself - Cultural engagement - engage in a	-1	Missing	000001
I_Q05aEEX			1	Never	000000
I_Q05aEEX			2	Less than once a mon	100000
I_Q05aEEX			3	Less than once a wee	010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
I_Q05aEEX			4	At least once a week	001000
I_Q05aEEX			5	Every day	000100
I_Q05aEEX			6	Valid skip	000010
I_Q05bEEX	7	About yourself - Cultural engagement - go to the m	-1	Missing	000001
I_Q05bEEX			1	Never	000000
I_Q05bEEX			2	Less than once a mon	100000
I_Q05bEEX			3	Less than once a wee	010000
I_Q05bEEX			4	At least once a week	001000
I_Q05bEEX			5	Every day	000100
I_Q05bEEX			6	Valid skip	000010
I_Q05cEEX	7	About yourself - Cultural engagement - particpate	-1	Missing	000001
I_Q05cEEX			1	Never	000000
I_Q05cEEX			2	Less than once a mon	100000
I_Q05cEEX			3	Less than once a wee	010000
I_Q05cEEX			4	At least once a week	001000
I_Q05cEEX			5	Every day	000100
I_Q05cEEX			6	Valid skip	000010
I_Q05dEEX	7	About yourself - Cultural engagement - visit a lib	-1	Missing	000001
I_Q05dEEX			1	Never	000000
I_Q05dEEX			2	Less than once a mon	100000
I_Q05dEEX			3	Less than once a wee	010000
I_Q05dEEX			4	At least once a week	001000
I_Q05dEEX			5	Every day	000100
I_Q05dEEX			6	Valid skip	000010
I_Q05eEEX	7	About yourself - Cultural engagement - spend time	-1	Missing	000001
I_Q05eEEX			1	Never	000000
I_Q05eEEX			2	Less than once a mon	100000
I_Q05eEEX			3	Less than once a wee	010000
I_Q05eEEX			4	At least once a week	001000
I_Q05eEEX			5	Every day	000100
I_Q05eEEX			6	Valid skip	000010
I_Q05hJPX	7	About yourself - Cultural engagement - Attend reli	-1	Missing	000001
I_Q05hJPX			1	Never	000000
I_Q05hJPX			2	Less than once a mon	100000
I_Q05hJPX			3	Less than once a wee	010000
I_Q05hJPX			4	At least once a week	001000
I_Q05hJPX			5	Every day	000100
I_Q05hJPX			6	Valid skip	000010
I_Q06c	7	I feel that I have a pretty good understanding of	-1	Missing	000001
I_Q06c			1	Strongly agree	000000
I_Q06c			2	Agree	100000
I_Q06c			3	Neither agree nor di	010000
I_Q06c			4	Disagree	001000
I_Q06c			5	Strongly disagree	000100
I_Q06c			6	Valid skip	000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
I_Q06dUSX1a	6	About yourself - Political efficacy - Information	-1	Missing	00001
I_Q06dUSX1a			1	A lot	00000
I_Q06dUSX1a			2	Some	10000
I_Q06dUSX1a			3	A little	01000
I_Q06dUSX1a			4	None	00100
I_Q06dUSX1a			6	Valid skip	00010
I_Q06dUSX1b	6	About yourself - Political efficacy - Information	-1	Missing	00001
I_Q06dUSX1b			1	A lot	00000
I_Q06dUSX1b			2	Some	10000
I_Q06dUSX1b			3	A little	01000
I_Q06dUSX1b			4	None	00100
I_Q06dUSX1b			6	Valid skip	00010
I_Q06dUSX1c	6	About yourself - Political efficacy - Information	-1	Missing	00001
I_Q06dUSX1c			1	A lot	00000
I_Q06dUSX1c			2	Some	10000
I_Q06dUSX1c			3	A little	01000
I_Q06dUSX1c			4	None	00100
I_Q06dUSX1c			6	Valid skip	00010
I_Q06dUSX1d	6	About yourself - Political efficacy - Information	-1	Missing	00001
I_Q06dUSX1d			1	A lot	00000
I_Q06dUSX1d			2	Some	10000
I_Q06dUSX1d			3	A little	01000
I_Q06dUSX1d			4	None	00100
I_Q06dUSX1d			6	Valid skip	00010
I_Q06dUSX1e	6	About yourself - Political efficacy - Information	-1	Missing	00001
I_Q06dUSX1e			1	A lot	00000
I_Q06dUSX1e			2	Some	10000
I_Q06dUSX1e			3	A little	01000
I_Q06dUSX1e			4	None	00100
I_Q06dUSX1e			6	Valid skip	00010
I_Q06dUSX1f	6	About yourself - Political efficacy - Information	-1	Missing	00001
I_Q06dUSX1f			1	A lot	00000
I_Q06dUSX1f			2	Some	10000
I_Q06dUSX1f			3	A little	01000
I_Q06dUSX1f			4	None	00100
I_Q06dUSX1f			6	Valid skip	00010
I_Q06dUSX1g	6	About yourself - Political efficacy - Information	-1	Missing	00001
I_Q06dUSX1g			1	A lot	00000
I_Q06dUSX1g			2	Some	10000
I_Q06dUSX1g			3	A little	01000
I_Q06dUSX1g			4	None	00100
I_Q06dUSX1g			6	Valid skip	00010
I_Q08USX1	4	About yourself - Health - Difficulty seeing print	-1	Missing	001
I_Q08USX1			1	Yes	000
I_Q08USX1			2	No	100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
I_Q08USX1			6	Valid skip	010
I_Q08USX2	4	About yourself - Health - Difficulty hearing conve	-1	Missing	001
I_Q08USX2			1	Yes	000
I_Q08USX2			2	No	100
I_Q08USX2			6	Valid skip	010
I_Q08USX3	4	About yourself - Health - Diagnosed learning disab	-1	Missing	001
I_Q08USX3			1	Yes	000
I_Q08USX3			2	No	100
I_Q08USX3			6	Valid skip	010
I_Q10a	4	About yourself - Disability - Longstanding illness	-1	Missing	001
I_Q10a			1	Yes	000
I_Q10a			2	No	100
I_Q10a			6	Valid skip	010
I_Q10aAUX	4	About yourself - Disability - Longstanding illness	-1	Missing	001
I_Q10aAUX			1	Yes	000
I_Q10aAUX			2	No	100
I_Q10aAUX			6	Valid skip	010
I_Q10aCA	4	About yourself - Disability - Longstanding illness	-1	Missing	001
I_Q10aCA			1	Yes	000
I_Q10aCA			2	No	100
I_Q10aCA			6	Valid skip	010
I_Q10b	5	About yourself - Disability - Limitiations because	-1	Missing	0001
I_Q10b			1	Severely limited	0000
I_Q10b			2	Limited but not seve	1000
I_Q10b			3	Not limited at all	0100
I_Q10b			6	Valid skip	0010
I_Q10bAUX	5	About yourself - Disability - Limitiations because	-1	Missing	0001
I_Q10bAUX			1	Severely limited	0000
I_Q10bAUX			2	Limited but not seve	1000
I_Q10bAUX			3	Not limited at all	0100
I_Q10bAUX			6	Valid skip	0010
I_Q10bCA	5	About yourself - Disability - Limitation because o	-1	Missing	0001
I_Q10bCA			1	Severely limited	0000
I_Q10bCA			2	Limited but not seve	1000
I_Q10bCA			3	Not limited at all	0100
I_Q10bCA			6	Valid skip	0010
I_Q10bUSX2a	6	About yourself - Health - Health information from	-1	Missing	00001
I_Q10bUSX2a			1	A lot	00000
I_Q10bUSX2a			2	Some	10000
I_Q10bUSX2a			3	A little	01000
I_Q10bUSX2a			4	None	00100
I_Q10bUSX2a			6	Valid skip	00010
I_Q10bUSX2b	6	About yourself - Health - Health information from	-1	Missing	00001
I_Q10bUSX2b			1	A lot	00000
I_Q10bUSX2b			2	Some	10000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
I_Q10bUSX2b			3	A little	01000
I_Q10bUSX2b			4	None	00100
I_Q10bUSX2b			6	Valid skip	00010
I_Q10bUSX2c	6	About yourself - Health - Health information from	-1	Missing	00001
I_Q10bUSX2c			1	A lot	00000
I_Q10bUSX2c			2	Some	10000
I_Q10bUSX2c			3	A little	01000
I_Q10bUSX2c			4	None	00100
I_Q10bUSX2c			6	Valid skip	00010
I_Q10bUSX2d	6	About yourself - Health - Health information from	-1	Missing	00001
I_Q10bUSX2d			1	A lot	00000
I_Q10bUSX2d			2	Some	10000
I_Q10bUSX2d			3	A little	01000
I_Q10bUSX2d			4	None	00100
I_Q10bUSX2d			6	Valid skip	00010
I_Q10bUSX2e	6	About yourself - Health - Health information from	-1	Missing	00001
I_Q10bUSX2e			1	A lot	00000
I_Q10bUSX2e			2	Some	10000
I_Q10bUSX2e			3	A little	01000
I_Q10bUSX2e			4	None	00100
I_Q10bUSX2e			6	Valid skip	00010
I_Q10bUSX2f	6	About yourself - Health - Health information from	-1	Missing	00001
I_Q10bUSX2f			1	A lot	00000
I_Q10bUSX2f			2	Some	10000
I_Q10bUSX2f			3	A little	01000
I_Q10bUSX2f			4	None	00100
I_Q10bUSX2f			6	Valid skip	00010
I_Q10bUSX2g	6	About yourself - Health - Health information from	-1	Missing	00001
I_Q10bUSX2g			1	A lot	00000
I_Q10bUSX2g			2	Some	10000
I_Q10bUSX2g			3	A little	01000
I_Q10bUSX2g			4	None	00100
I_Q10bUSX2g			6	Valid skip	00010
I_Q10bUSX2h	6	About yourself - Health - Health information from	-1	Missing	00001
I_Q10bUSX2h			1	A lot	00000
I_Q10bUSX2h			2	Some	10000
I_Q10bUSX2h			3	A little	01000
I_Q10bUSX2h			4	None	00100
I_Q10bUSX2h			6	Valid skip	00010
I_Q10bUSX3a	4	About yourself - Health - Flu shot in past year	-1	Missing	001
I_Q10bUSX3a			1	Yes	000
I_Q10bUSX3a			2	No	100
I_Q10bUSX3a			6	Valid skip	010
I_Q10bUSX3b	4	About yourself - Health - Mammogram in past year	-1	Missing	001
I_Q10bUSX3b			1	Yes	000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
I_Q10bUSX3b			2	No	100
I_Q10bUSX3b			6	Valid skip	010
I_Q10bUSX3c	4	About yourself - Health - Pap smear in past year	-1	Missing	001
I_Q10bUSX3c			1	Yes	000
I_Q10bUSX3c			2	No	100
I_Q10bUSX3c			6	Valid skip	010
I_Q10bUSX3d	4	About yourself - Health - Screen for colon cancer	-1	Missing	001
I_Q10bUSX3d			1	Yes	000
I_Q10bUSX3d			2	No	100
I_Q10bUSX3d			6	Valid skip	010
I_Q10bUSX3e	4	About yourself - Health - Vision check in past yea	-1	Missing	001
I_Q10bUSX3e			1	Yes	000
I_Q10bUSX3e			2	No	100
I_Q10bUSX3e			6	Valid skip	010
I_Q10bUSX3f	4	About yourself - Health - Screen for prostate canc	-1	Missing	001
I_Q10bUSX3f			1	Yes	000
I_Q10bUSX3f			2	No	100
I_Q10bUSX3f			6	Valid skip	010
I_Q10bUSX3g	4	About yourself - Health - Screen for osteoporosis	-1	Missing	001
I_Q10bUSX3g			1	Yes	000
I_Q10bUSX3g			2	No	100
I_Q10bUSX3g			6	Valid skip	010
I_Q10bUSX3h	4	About yourself - Health - Seen dentist in past yea	-1	Missing	001
I_Q10bUSX3h			1	Yes	000
I_Q10bUSX3h			2	No	100
I_Q10bUSX3h			6	Valid skip	010
I_Q10UKX	5	About yourself - Disability - Day-to-day activitie	-1	Missing	0001
I_Q10UKX			1	Yes, limited a lot	0000
I_Q10UKX			2	Yes, limited a littl	1000
I_Q10UKX			3	No	0100
I_Q10UKX			6	Valid skip	0010
J_N05a2DK	4	Did the respondent mention more than 1 language?	-1	Missing	001
J_N05a2DK			1	Yes	000
J_N05a2DK			2	No	100
J_N05a2DK			96	Valid skip	010
J_N05bDEX1	4	Background - More than one language spoken at home	-1	Missing	001
J_N05bDEX1			1	Yes	000
J_N05bDEX1			2	No	100
J_N05bDEX1			6	Valid skip	010
J_N05bDEX2	4	Background - More than one language spoken at age	-1	Missing	001
J_N05bDEX2			1	Yes	000
J_N05bDEX2			2	No	100
J_N05bDEX2			6	Valid skip	010
J_Q01AU	8	Background - People in household AU	-1	Missing	0000001
J_Q01AU			1	1	0000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q01AU			2	2	1000000
J_Q01AU			3	3	0100000
J_Q01AU			4	4	0010000
J_Q01AU			5	5	0001000
J_Q01AU			6	6 or more	0000100
J_Q01AU			96	Valid skip	0000010
J_Q02aUK	12	Background - Living with spouse or partner	-1	Missing	00000000001
J_Q02aUK			1	single, that is neve	00000000000
J_Q02aUK			2	married and living w	10000000000
J_Q02aUK			3	living with someone	01000000000
J_Q02aUK			4	a civil partner in a	00100000000
J_Q02aUK			5	married and separate	00010000000
J_Q02aUK			6	divorced	00001000000
J_Q02aUK			7	widowed	00000100000
J_Q02aUK			8	Spontaneous only - L	00000010000
J_Q02aUK			9	Spontaneous only - C	00000001000
J_Q02aUK			10	Spontaneous only - S	00000000100
J_Q02aUK			96	Valid skip	00000000010
J_Q02bCZ	15	Background - Highest education level partner has e	-1	Missing	00000000000001
J_Q02bCZ			1	ISCED 1	00000000000000
J_Q02bCZ			2	ISCED 2	10000000000000
J_Q02bCZ			3	ISCED 3C shorter tha	01000000000000
J_Q02bCZ			4	ISCED 3C 2 years or	00100000000000
J_Q02bCZ			5	ISCED 3A-B	00010000000000
J_Q02bCZ			6	ISCED 3 (without dis	00001000000000
J_Q02bCZ			7	ISCED 4C	00000100000000
J_Q02bCZ			8	ISCED 4A-B	00000010000000
J_Q02bCZ			9	ISCED 4 (without dis	00000001000000
J_Q02bCZ			10	ISCED 5B	00000000100000
J_Q02bCZ			11	ISCED 5A, bachelor d	00000000010000
J_Q02bCZ			12	ISCED 5A, master deg	00000000001000
J_Q02bCZ			13	ISCED 6	00000000000100
J_Q02bCZ			96	Valid skip	00000000000010
J_Q02bFR	16	Background - Highest education level partner has e	-1	Missing	000000000000001
J_Q02bFR			1	No formal qualificat	000000000000000
J_Q02bFR			2	ISCED 1	100000000000000
J_Q02bFR			3	ISCED 2	010000000000000
J_Q02bFR			4	ISCED 3C shorter tha	001000000000000
J_Q02bFR			5	ISCED 3C 2 years or	000100000000000
J_Q02bFR			6	ISCED 3A-B	000010000000000
J_Q02bFR			7	ISCED 3 (without dis	000001000000000
J_Q02bFR			8	ISCED 4C	000000100000000
J_Q02bFR			9	ISCED 4A-B	000000010000000
J_Q02bFR			10	ISCED 4 (without dis	000000001000000
J_Q02bFR			11	ISCED 5B	000000000100000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q02bFR			12	ISCED 5A, bachelor d	000000000010000
J_Q02bFR			13	ISCED 5A, master deg	000000000001000
J_Q02bFR			14	ISCED 6	000000000000100
J_Q02bFR			96	Valid skip	000000000000010
J_Q02cCZ	11	Background - Work situation of spouse of partner	-1	Missing	0000000001
J_Q02cCZ			1	Full-time employed (	0000000000
J_Q02cCZ			2	Part-time employed (	1000000000
J_Q02cCZ			3	Unemployed	0100000000
J_Q02cCZ			4	Pupil, student	0010000000
J_Q02cCZ			5	Apprentice, internsh	0001000000
J_Q02cCZ			6	In retirement or ear	0000100000
J_Q02cCZ			7	Permanently disabled	0000010000
J_Q02cCZ			8	Fulfilling domestic	0000001000
J_Q02cCZ			9	Other	0000000100
J_Q02cCZ			96	Valid skip	0000000010
J_Q02cIE	11	Background - Work situation of spouse or partner	-1	Missing	0000000001
J_Q02cIE			1	Full-time employed (	0000000000
J_Q02cIE			2	Part-time employed (	1000000000
J_Q02cIE			3	Unemployed	0100000000
J_Q02cIE			4	Pupil, student	0010000000
J_Q02cIE			5	Apprentice, internsh	0001000000
J_Q02cIE			6	In retirement or ear	0000100000
J_Q02cIE			7	Permanently disabled	0000010000
J_Q02cIE			8	Fulfilling domestic	0000001000
J_Q02cIE			9	Other	0000000100
J_Q02cIE			96	Valid skip	0000000010
J_Q02cNL	11	Background - Work situation of spouse or partner	-1	Missing	0000000001
J_Q02cNL			1	Full-time employed (	0000000000
J_Q02cNL			2	Part-time employed (	1000000000
J_Q02cNL			3	Unemployed	0100000000
J_Q02cNL			4	Pupil, student	0010000000
J_Q02cNL			5	Apprentice, internsh	0001000000
J_Q02cNL			6	In retirement or ear	0000100000
J_Q02cNL			7	Permanently disabled	0000010000
J_Q02cNL			8	Fulfilling domestic	0000001000
J_Q02cNL			9	Other	0000000100
J_Q02cNL			96	Valid skip	0000000010
J_Q03aAU	4	Background - Children	-1	Missing	001
J_Q03aAU			1	Yes	000
J_Q03aAU			2	No	100
J_Q03aAU			6	Valid skip	010
J_Q03bAUa	6	Background - Number of children (AUS)	-1	Missing	00001
J_Q03bAUa			1	1	00000
J_Q03bAUa			2	2	10000
J_Q03bAUa			3	3	01000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q03bAUa			4	4 or more	00100
J_Q03bAUa			96	Valid skip	00010
J_Q03cAUa	22	Background - Age of the child (AUS) - grouped	-1	Missing	000000000000000000001
J_Q03cAUa			1	0-4 years	000000000000000000000
J_Q03cAUa			2	5-9 years	100000000000000000000
J_Q03cAUa			3	10-14 years	010000000000000000000
J_Q03cAUa			4	15 years	001000000000000000000
J_Q03cAUa			5	16 years	000100000000000000000
J_Q03cAUa			6	17 years	000010000000000000000
J_Q03cAUa			7	18 years	000001000000000000000
J_Q03cAUa			8	19 years	000000100000000000000
J_Q03cAUa			9	20 years	000000010000000000000
J_Q03cAUa			10	21 years	000000001000000000000
J_Q03cAUa			11	22 years	000000000100000000000
J_Q03cAUa			12	23 years	000000000010000000000
J_Q03cAUa			13	24 years	000000000001000000000
J_Q03cAUa			14	25-29 years	000000000000100000000
J_Q03cAUa			15	30-34 years	000000000000010000000
J_Q03cAUa			16	35-39 years	000000000000001000000
J_Q03cAUa			17	40-44 years	000000000000000100000
J_Q03cAUa			18	45-49 years	000000000000000010000
J_Q03cAUa			19	50-54 years	000000000000000001000
J_Q03cAUa			20	55 years and over	000000000000000000100
J_Q03cAUa			96	Valid skip	000000000000000000010
J_Q03d1AUa	22	Background - Age of the youngest child (AUS) - gro	-1	Missing	000000000000000000001
J_Q03d1AUa			1	0-4 years	000000000000000000000
J_Q03d1AUa			2	5-9 years	100000000000000000000
J_Q03d1AUa			3	10-14 years	010000000000000000000
J_Q03d1AUa			4	15 years	001000000000000000000
J_Q03d1AUa			5	16 years	000100000000000000000
J_Q03d1AUa			6	17 years	000010000000000000000
J_Q03d1AUa			7	18 years	000001000000000000000
J_Q03d1AUa			8	19 years	000000100000000000000
J_Q03d1AUa			9	20 years	000000010000000000000
J_Q03d1AUa			10	21 years	000000001000000000000
J_Q03d1AUa			11	22 years	000000000100000000000
J_Q03d1AUa			12	23 years	000000000010000000000
J_Q03d1AUa			13	24 years	000000000001000000000
J_Q03d1AUa			14	25-29 years	000000000000100000000
J_Q03d1AUa			15	30-34 years	000000000000010000000
J_Q03d1AUa			16	35-39 years	000000000000001000000
J_Q03d1AUa			17	40-44 years	000000000000000100000
J_Q03d1AUa			18	45-49 years	000000000000000010000
J_Q03d1AUa			19	50-54 years	000000000000000001000
J_Q03d1AUa			20	55 years and over	000000000000000000100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q03d1AUa			96	Valid skip	000000000000000000010
J_Q03d2AUa	22	Background - Age of the oldest child (AUS) - group	-1	Missing	000000000000000000001
J_Q03d2AUa			1	0-4 years	000000000000000000000
J_Q03d2AUa			2	5-9 years	100000000000000000000
J_Q03d2AUa			3	10-14 years	010000000000000000000
J_Q03d2AUa			4	15 years	001000000000000000000
J_Q03d2AUa			5	16 years	000100000000000000000
J_Q03d2AUa			6	17 years	000010000000000000000
J_Q03d2AUa			7	18 years	000001000000000000000
J_Q03d2AUa			8	19 years	000000100000000000000
J_Q03d2AUa			9	20 years	000000010000000000000
J_Q03d2AUa			10	21 years	000000001000000000000
J_Q03d2AUa			11	22 years	000000000100000000000
J_Q03d2AUa			12	23 years	000000000010000000000
J_Q03d2AUa			13	24 years	000000000001000000000
J_Q03d2AUa			14	25-29 years	000000000000100000000
J_Q03d2AUa			15	30-34 years	000000000000010000000
J_Q03d2AUa			16	35-39 years	000000000000001000000
J_Q03d2AUa			17	40-44 years	000000000000000100000
J_Q03d2AUa			18	45-49 years	000000000000000010000
J_Q03d2AUa			19	50-54 years	000000000000000001000
J_Q03d2AUa			20	55 years and over	000000000000000000100
J_Q03d2AUa			96	Valid skip	000000000000000000010
J_Q03UKX	4	Background - Caring for live-in elderly/long-term	-1	Missing	001
J_Q03UKX			1	Yes	000
J_Q03UKX			2	No	100
J_Q03UKX			6	Valid skip	010
J_Q04aAU	4	Background - Born in Australia	-1	Missing	001
J_Q04aAU			1	Yes	000
J_Q04aAU			2	No	100
J_Q04aAU			6	Valid skip	010
J_Q04aRU	4	Background - Born in country	-1	Missing	001
J_Q04aRU			1	Yes	000
J_Q04aRU			2	No	100
J_Q04aRU			6	Valid skip	010
J_Q04bAT	15	Background - Country of birth - NATIONAL	-1	Missing	00000000000001
J_Q04bAT			1	Bosnia and Herzegovi	00000000000000
J_Q04bAT			2	Germany	10000000000000
J_Q04bAT			3	Italy	01000000000000
J_Q04bAT			4	Croatia	00100000000000
J_Q04bAT			5	Montenegro	00010000000000
J_Q04bAT			6	Poland	00001000000000
J_Q04bAT			7	Rumania	00000100000000
J_Q04bAT			8	Russia	00000010000000
J_Q04bAT			9	Serbia	00000001000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q04bAT			10	Czech Republic	00000000100000
J_Q04bAT			11	Turkey	00000000010000
J_Q04bAT			12	Hungary	00000000001000
J_Q04bAT			13	Other country	00000000000100
J_Q04bAT			96	Valid skip	00000000000010
J_Q04bAU	12	Background - Country of birth (AUS)	-1	Missing	00000000001
J_Q04bAU			1	England	00000000000
J_Q04bAU			2	New Zealand	10000000000
J_Q04bAU			3	Italy	01000000000
J_Q04bAU			4	Viet Nam	00100000000
J_Q04bAU			5	India	00010000000
J_Q04bAU			6	Scotland	00001000000
J_Q04bAU			7	Philippines	00000100000
J_Q04bAU			8	Greece	00000010000
J_Q04bAU			9	Germany	00000001000
J_Q04bAU			10	Other	00000000100
J_Q04bAU			96	Valid skip	00000000010
J_Q04bAUa	4	Background - Country of birth (AUS)	-1	Missing	001
J_Q04bAUa			1	Main English speakin	000
J_Q04bAUa			2	Other countries	100
J_Q04bAUa			96	Valid skip	010
J_Q04bBE	12	Background - Country of birth	-1	Missing	00000000001
J_Q04bBE			1	The Netherlands	00000000000
J_Q04bBE			2	Italy	10000000000
J_Q04bBE			3	France	01000000000
J_Q04bBE			4	Germany	00100000000
J_Q04bBE			5	Spain	00010000000
J_Q04bBE			6	Morocco	00001000000
J_Q04bBE			7	Turkey	00000100000
J_Q04bBE			8	Poland	00000010000
J_Q04bBE			9	Former Yugoslavia	00000001000
J_Q04bBE			10	Other country	00000000100
J_Q04bBE			96	Valid skip	00000000010
J_Q04bca2	12	Background - Country of birth	-1	Missing	00000000001
J_Q04bca2			1	China (People's Repu	00000000000
J_Q04bca2			2	Germany	10000000000
J_Q04bca2			3	Hong Kong	01000000000
J_Q04bca2			4	India	00100000000
J_Q04bca2			5	Italy	00010000000
J_Q04bca2			6	Jamaica	00001000000
J_Q04bca2			7	Philippines	00000100000
J_Q04bca2			8	United Kingdom (e.g.	00000010000
J_Q04bca2			9	United States	00000001000
J_Q04bca2			10	Other - specify	00000000100
J_Q04bca2			96	Valid skip	00000000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q04bca3	4	Background - Canadian by birth, naturalization, la	-1	Missing	001
J_Q04bca3			1	Yes	000
J_Q04bca3			2	No	100
J_Q04bca3			6	Valid skip	010
J_Q04bca4	6	Background - Immigration programs	-1	Missing	00001
J_Q04bca4			1	N the refugee progra	00000
J_Q04bca4			2	N the program of re-	10000
J_Q04bca4			3	... the points syste	01000
J_Q04bca4			4	... or other?	00100
J_Q04bca4			6	Valid skip	00010
J_Q04bca7	4	Background - First came to Canada as a refugee	-1	Missing	001
J_Q04bca7			1	Yes	000
J_Q04bca7			2	No	100
J_Q04bca7			6	Valid skip	010
J_Q04bCY	8	Background - Country of birth	-1	Missing	0000001
J_Q04bCY			1	Greece	0000000
J_Q04bCY			2	United Kingdom	1000000
J_Q04bCY			3	Russian Federation	0100000
J_Q04bCY			4	Bulgaria	0010000
J_Q04bCY			5	Georgia	0001000
J_Q04bCY			6	Other country	0000100
J_Q04bCY			96	Valid skip	0000010
J_Q04bCZ	9	Background - Country of birth	-1	Missing	00000001
J_Q04bCZ			1	Country 1	00000000
J_Q04bCZ			2	Country 2	10000000
J_Q04bCZ			3	Country 3	01000000
J_Q04bCZ			4	Country 4	00100000
J_Q04bCZ			5	Country 5	00010000
J_Q04bCZ			6	Country 6	00001000
J_Q04bCZ			7	Other country	00000100
J_Q04bCZ			96	Valid skip	00000010
J_Q04bDE	11	Background - Country of birth	-1	Missing	0000000001
J_Q04bDE			1	Turkey	0000000000
J_Q04bDE			2	Italy	1000000000
J_Q04bDE			3	Poland	0100000000
J_Q04bDE			4	Greece	0010000000
J_Q04bDE			5	Serbia	0001000000
J_Q04bDE			6	Croatia	0000100000
J_Q04bDE			7	Russian Federation	0000010000
J_Q04bDE			8	Bosnia and Herzegovi	0000001000
J_Q04bDE			9	Another country	0000000100
J_Q04bDE			96	Valid skip	0000000010
J_Q04bDK	9	Background - Country of birth	-1	Missing	00000001
J_Q04bDK			1	Turkey	00000000
J_Q04bDK			2	Germany	10000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q04bDK			3	Poland	01000000
J_Q04bDK			4	Iraq	00100000
J_Q04bDK			5	Bosnia-Herzegovinia	00010000
J_Q04bDK			6	Norway	00001000
J_Q04bDK			7	Other country	00000100
J_Q04bDK			96	Valid skip	00000010
J_Q04bEE	8	Background - Country of birth	-1	Missing	0000001
J_Q04bEE			1	Russia	0000000
J_Q04bEE			2	Ukraine	1000000
J_Q04bEE			3	Belarus	0100000
J_Q04bEE			4	Latvia	0010000
J_Q04bEE			5	Finland	0001000
J_Q04bEE			6	Other country	0000100
J_Q04bEE			96	Valid skip	0000010
J_Q04bES	13	Background - Country of birth	-1	Missing	000000000001
J_Q04bES			1	Alemania	000000000000
J_Q04bES			2	Argentina	100000000000
J_Q04bES			3	Colombia	010000000000
J_Q04bES			4	Ecuador	001000000000
J_Q04bES			5	Marruecos	000100000000
J_Q04bES			6	Marruecos	000010000000
J_Q04bES			7	Reino Unido	000001000000
J_Q04bES			8	Reinblica Dominicana	000000100000
J_Q04bES			9	Reinbla	000000010000
J_Q04bES			10	Venezuela	000000001000
J_Q04bES			11	Venezuels	000000000100
J_Q04bES			96	Valid skip	000000000010
J_Q04bFI	7	Background - Country of birth	-1	Missing	000001
J_Q04bFI			1	Sweden	000000
J_Q04bFI			2	Russia	100000
J_Q04bFI			3	Former Soviet Union	010000
J_Q04bFI			4	Estonia	001000
J_Q04bFI			5	Other country	000100
J_Q04bFI			96	Valid skip	000010
J_Q04bFR	12	Background - Country of birth	-1	Missing	00000000001
J_Q04bFR			1	Algeria	00000000000
J_Q04bFR			2	Germany	10000000000
J_Q04bFR			3	Spain	01000000000
J_Q04bFR			4	Italy	00100000000
J_Q04bFR			5	Morocco	00010000000
J_Q04bFR			6	Portugal	00001000000
J_Q04bFR			7	United Kingdom	00000100000
J_Q04bFR			8	Tunisia	00000010000
J_Q04bFR			9	Turkey	00000001000
J_Q04bFR			10	Other countries	00000000100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q04bFR			96	Valid skip	00000000010
J_Q04bIE	10	Background - Country of birth	-1	Missing	000000001
J_Q04bIE			1	Poland	000000000
J_Q04bIE			2	United Kingdom	100000000
J_Q04bIE			3	Lithuania	010000000
J_Q04bIE			4	Latvia	001000000
J_Q04bIE			5	Germany	000100000
J_Q04bIE			6	Romania	000010000
J_Q04bIE			7	Northern Ireland	000001000
J_Q04bIE			8	Other country	000000100
J_Q04bIE			96	Valid skip	000000010
J_Q04bIT	18	Background - Country of birth	-1	Missing	00000000000000001
J_Q04bIT			1	Albania	00000000000000000
J_Q04bIT			2	China	10000000000000000
J_Q04bIT			3	Ecuador	01000000000000000
J_Q04bIT			4	Philippines	00100000000000000
J_Q04bIT			5	France	00010000000000000
J_Q04bIT			6	Germany	00001000000000000
J_Q04bIT			7	Morocco	00000100000000000
J_Q04bIT			8	Peru	00000010000000000
J_Q04bIT			9	Poland	00000001000000000
J_Q04bIT			10	United Kingdom	00000000100000000
J_Q04bIT			11	Romania	00000000010000000
J_Q04bIT			12	Spain	00000000001000000
J_Q04bIT			13	United States of Ame	00000000000100000
J_Q04bIT			14	Tunisia	00000000000010000
J_Q04bIT			15	Ukraina	00000000000001000
J_Q04bIT			16	Other	00000000000000100
J_Q04bIT			96	Valid skip	00000000000000010
J_Q04bJP	12	Background - Country of birth	-1	Missing	00000000001
J_Q04bJP			1	USA	00000000000
J_Q04bJP			2	Canada	10000000000
J_Q04bJP			3	UK	01000000000
J_Q04bJP			4	Australia	00100000000
J_Q04bJP			5	New Zealand	00010000000
J_Q04bJP			6	Republic of Korea	00001000000
J_Q04bJP			7	China	00000100000
J_Q04bJP			8	Germany	00000010000
J_Q04bJP			9	France	00000001000
J_Q04bJP			10	Other country	00000000100
J_Q04bJP			96	Valid skip	00000000010
J_Q04bKO	9	KO_Background - Country of birth	-1	Missing	00000001
J_Q04bKO			1	China	00000000
J_Q04bKO			2	United States	10000000
J_Q04bKO			3	Vietnam	01000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q04bKO			4	Philippines	00100000
J_Q04bKO			5	Thailand	00010000
J_Q04bKO			6	Japan	00001000
J_Q04bKO			7	Other country	00000100
J_Q04bKO			96	Valid skip	00000010
J_Q04bNL	9	Background - Country of birth	-1	Missing	00000001
J_Q04bNL			1	Marocco	00000000
J_Q04bNL			2	Turkey	10000000
J_Q04bNL			3	Surinam	01000000
J_Q04bNL			4	Dutch Antillen	00100000
J_Q04bNL			5	Germany	00010000
J_Q04bNL			6	Belgium	00001000
J_Q04bNL			7	Other country	00000100
J_Q04bNL			96	Valid skip	00000010
J_Q04bNO	10	Background - Country of birth	-1	Missing	000000001
J_Q04bNO			1	Polan	000000000
J_Q04bNO			2	Sweden	100000000
J_Q04bNO			3	Pakistan	010000000
J_Q04bNO			4	Iraq	001000000
J_Q04bNO			5	Iran	000100000
J_Q04bNO			6	Somalia	000010000
J_Q04bNO			7	USA	000001000
J_Q04bNO			8	Other country	000000100
J_Q04bNO			186	Valid skip	000000010
J_Q04bPL	15	Background - Country of birth	-1	Missing	00000000000001
J_Q04bPL			1	Belarus	00000000000000
J_Q04bPL			2	Czech Republic	10000000000000
J_Q04bPL			3	England	01000000000000
J_Q04bPL			4	France	00100000000000
J_Q04bPL			5	Germany	00010000000000
J_Q04bPL			6	Lithuania	00001000000000
J_Q04bPL			7	Netherlands	00000100000000
J_Q04bPL			8	Poland	00000010000000
J_Q04bPL			9	Russia	00000001000000
J_Q04bPL			10	Slovakia	00000000100000
J_Q04bPL			11	Ukraine	00000000010000
J_Q04bPL			12	United States of Ame	00000000001000
J_Q04bPL			13	Other country	00000000000100
J_Q04bPL			96	Valid skip	00000000000010
J_Q04bRU	8	Background - Country of birth	-1	Missing	0000001
J_Q04bRU			1	Country 1	0000000
J_Q04bRU			2	Country 2	1000000
J_Q04bRU			3	Country 3	0100000
J_Q04bRU			4	Country 4	0010000
J_Q04bRU			5	Country 5	0001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q04bRU			6	Other country	0000100
J_Q04bRU			96	Valid skip	0000010
J_Q04bSE	15	Background - Country of birth	-1	Missing	00000000000001
J_Q04bSE			1	Finland	00000000000000
J_Q04bSE			2	Irak	10000000000000
J_Q04bSE			3	Serbien	01000000000000
J_Q04bSE			4	Iran	00100000000000
J_Q04bSE			5	Polen	00010000000000
J_Q04bSE			6	Bosnien-Hercegovina	00001000000000
J_Q04bSE			7	Turkiet	00000100000000
J_Q04bSE			8	Danmark	00000010000000
J_Q04bSE			9	Norge	00000001000000
J_Q04bSE			10	Chile	00000000100000
J_Q04bSE			11	Tyskland	00000000010000
J_Q04bSE			12	Kroatien	00000000001000
J_Q04bSE			13	Annat land var god a	00000000000100
J_Q04bSE			96	Valid skip	00000000000010
J_Q04bSK	9	Background - Country of birth	-1	Missing	00000001
J_Q04bSK			1	Czech republic	00000000
J_Q04bSK			2	Hungary	10000000
J_Q04bSK			3	Austria	01000000
J_Q04bSK			4	Poland	00100000
J_Q04bSK			5	Germany	00010000
J_Q04bSK			6	Ukraine	00001000
J_Q04bSK			7	other country	00000100
J_Q04bSK			96	Valid skip	00000010
J_Q04bUK	16	Background - Country of birth	-1	Missing	000000000000001
J_Q04bUK			1	India	000000000000000
J_Q04bUK			2	Poland	100000000000000
J_Q04bUK			3	Pakistan	010000000000000
J_Q04bUK			4	Germany	001000000000000
J_Q04bUK			5	South Africa	000100000000000
J_Q04bUK			6	Bangladesh	000010000000000
J_Q04bUK			7	Nigeria	000001000000000
J_Q04bUK			8	Kenya	000000100000000
J_Q04bUK			9	United States	000000010000000
J_Q04bUK			10	Phillippines	000000001000000
J_Q04bUK			11	France	000000000100000
J_Q04bUK			12	Australia	000000000010000
J_Q04bUK			13	Republic of Ireland	000000000001000
J_Q04bUK			14	Other Country	000000000000100
J_Q04bUK			96	Valid skip	000000000000010
J_Q04bUS	9	Background - Country of birth	-1	Missing	00000001
J_Q04bUS			1	Mexico	00000000
J_Q04bUS			2	China	10000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q04bUS			3	Phillipines	01000000
J_Q04bUS			4	India	00100000
J_Q04bUS			5	Russia	00010000
J_Q04bUS			6	Colombia	00001000
J_Q04bUS			7	Other country	00000100
J_Q04bUS			96	Valid skip	00000010
J_Q04c2ATX	9	Background - Citizenship - NATIONAL	-1	Missing	00000001
J_Q04c2ATX			1	Austria	00000000
J_Q04c2ATX			2	Germany	10000000
J_Q04c2ATX			3	Serbia	01000000
J_Q04c2ATX			4	Turkey	00100000
J_Q04c2ATX			5	Bosnia and Herzegovi	00010000
J_Q04c2ATX			6	Croatia	00001000
J_Q04c2ATX			7	Other country	00000100
J_Q04c2ATX			96	Valid skip	00000010
J_Q04c2AUa	5	Background - Year of immigration (AUS)	-1	Missing	0001
J_Q04c2AUa			1	Arrived 1991 or befo	0000
J_Q04c2AUa			2	Arrived 1992-2001	1000
J_Q04c2AUa			3	Arrived 2002-2012	0100
J_Q04c2AUa			9996	Valid skip	0010
J_Q04c2DEX1	4	Background - Citizenship - German	-1	Missing	001
J_Q04c2DEX1			1	Yes	000
J_Q04c2DEX1			2	No	100
J_Q04c2DEX1			6	Valid skip	010
J_Q04c2DEX2	4	Background - Citizenship - Additional to German	-1	Missing	001
J_Q04c2DEX2			1	Yes	000
J_Q04c2DEX2			2	No	100
J_Q04c2DEX2			6	Valid skip	010
J_Q04c2DEX3	13	Background - Citizenship - (Second) Citizenship -	-1	Missing	000000000001
J_Q04c2DEX3			1	Turkey	000000000000
J_Q04c2DEX3			2	Italy	100000000000
J_Q04c2DEX3			3	Poland	010000000000
J_Q04c2DEX3			4	Greece	001000000000
J_Q04c2DEX3			5	Serbia	000100000000
J_Q04c2DEX3			6	Croatia	000010000000
J_Q04c2DEX3			7	Bosnia and Herzegovi	000001000000
J_Q04c2DEX3			8	Macedonia	000000100000
J_Q04c2DEX3			9	Slovenia	000000010000
J_Q04c2DEX3			10	Russian Federation	000000001000
J_Q04c2DEX3			11	Another citizenship	000000000100
J_Q04c2DEX3			96	Valid skip	000000000010
J_Q04c2DEX4	4	Background - Residence before German reunification	-1	Missing	001
J_Q04c2DEX4			1	In the GDR	000
J_Q04c2DEX4			2	In the Federal Repub	100
J_Q04c2DEX4			6	Valid skip	010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q04c2EEX	4	Background - Lived in another country	-1	Missing	001
J_Q04c2EEX			1	Yes	000
J_Q04c2EEX			2	No	100
J_Q04c2EEX			6	Valid skip	010
J_Q04dUSX1a	4	Background - Hispanic	-1	Missing	001
J_Q04dUSX1a			1	Yes	000
J_Q04dUSX1a			2	No	100
J_Q04dUSX1a			6	Valid skip	010
J_Q04dUSX1b_01	4	Background - Hispanic origin - Mexican	-1	Missing	001
J_Q04dUSX1b_01			1	Marked	000
J_Q04dUSX1b_01			2	Not marked	100
J_Q04dUSX1b_01			6	Valid skip	010
J_Q04dUSX1b_02	4	Background - Hispanic origin - Puerto Rican	-1	Missing	001
J_Q04dUSX1b_02			1	Marked	000
J_Q04dUSX1b_02			2	Not marked	100
J_Q04dUSX1b_02			6	Valid skip	010
J_Q04dUSX1b_03	4	Background - Hispanic origin - Cuban	-1	Missing	001
J_Q04dUSX1b_03			1	Marked	000
J_Q04dUSX1b_03			2	Not marked	100
J_Q04dUSX1b_03			6	Valid skip	010
J_Q04dUSX1b_04	4	Background - Hispanic origin - Central/South Ameri	-1	Missing	001
J_Q04dUSX1b_04			1	Marked	000
J_Q04dUSX1b_04			2	Not marked	100
J_Q04dUSX1b_04			6	Valid skip	010
J_Q04dUSX1b_05	4	Background - Hispanic origin - Other	-1	Missing	001
J_Q04dUSX1b_05			1	Marked	000
J_Q04dUSX1b_05			2	Not marked	100
J_Q04dUSX1b_05			6	Valid skip	010
J_Q04dUSX2_01	4	Background - Race - White	-1	Missing	001
J_Q04dUSX2_01			1	Marked	000
J_Q04dUSX2_01			2	Not marked	100
J_Q04dUSX2_01			6	Valid skip	010
J_Q04dUSX2_02	4	Background - Race - Black	-1	Missing	001
J_Q04dUSX2_02			1	Marked	000
J_Q04dUSX2_02			2	Not marked	100
J_Q04dUSX2_02			6	Valid skip	010
J_Q04dUSX2_03	4	Background - Race - Asian	-1	Missing	001
J_Q04dUSX2_03			1	Marked	000
J_Q04dUSX2_03			2	Not marked	100
J_Q04dUSX2_03			6	Valid skip	010
J_Q04dUSX2_04	4	Background - Race - American Indian	-1	Missing	001
J_Q04dUSX2_04			1	Marked	000
J_Q04dUSX2_04			2	Not marked	100
J_Q04dUSX2_04			6	Valid skip	010
J_Q04dUSX2_05	4	Background - Race - Native Hawaiian	-1	Missing	001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q04dUSX2_05			1	Marked	000
J_Q04dUSX2_05			2	Not marked	100
J_Q04dUSX2_05			6	Valid skip	010
J_Q04eca1	4	Background - English/French language training	-1	Missing	001
J_Q04eca1			1	Yes	000
J_Q04eca1			2	No	100
J_Q04eca1			6	Valid skip	010
J_Q04eca2	7	Background - Planning to take English/French langu	-1	Missing	000001
J_Q04eca2			1	Yes, within the next	000000
J_Q04eca2			2	Yes, within the next	100000
J_Q04eca2			3	Yes, within the next	010000
J_Q04eca2			4	Yes, but not sure wh	001000
J_Q04eca2			5	No	000100
J_Q04eca2			6	Valid skip	000010
J_Q04fca1	4	Background - Aboriginal person	-1	Missing	001
J_Q04fca1			1	Yes	000
J_Q04fca1			2	No	100
J_Q04fca1			6	Valid skip	010
J_Q04fca2_01	4	Background - Aboriginal person	-1	Missing	001
J_Q04fca2_01			1	Marked	000
J_Q04fca2_01			2	Not marked	100
J_Q04fca2_01			6	Valid skip	010
J_Q04fca2_02	4	Background - Aboriginal person	-1	Missing	001
J_Q04fca2_02			1	Marked	000
J_Q04fca2_02			2	Not marked	100
J_Q04fca2_02			6	Valid skip	010
J_Q04fca2_03	4	Background - Aboriginal person	-1	Missing	001
J_Q04fca2_03			1	Marked	000
J_Q04fca2_03			2	Not marked	100
J_Q04fca2_03			6	Valid skip	010
J_Q04fca3	4	Background - Aboriginal person - Status Indian (Re	-1	Missing	001
J_Q04fca3			1	Yes, Status Indian (	000
J_Q04fca3			2	No	100
J_Q04fca3			6	Valid skip	010
J_Q04fca4	4	Background - Aboriginal person - Member of a First	-1	Missing	001
J_Q04fca4			1	Yes	000
J_Q04fca4			2	No	100
J_Q04fca4			6	Valid skip	010
J_Q04UKX1	7	Background - Ethnic group - white/mixed/asian/blac	-1	Missing	000001
J_Q04UKX1			1	White	000000
J_Q04UKX1			2	Mixed race	100000
J_Q04UKX1			3	Asian or Asian Briti	010000
J_Q04UKX1			4	Black or Black Briti	001000
J_Q04UKX1			5	Other ethnic group	000100
J_Q04UKX1			96	Valid skip	000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q04UKX10	4	Background - Ethnic group - other ethnic (Scot)	-1	Missing	001
J_Q04UKX10			1	Arab	000
J_Q04UKX10			2	Any other	100
J_Q04UKX10			96	Valid skip	010
J_Q04UKX2	9	Background - Ethnic group - UK english	-1	Missing	00000001
J_Q04UKX2			1	English	00000000
J_Q04UKX2			2	Scottish	10000000
J_Q04UKX2			3	Welsh	01000000
J_Q04UKX2			4	Northern Irish	00100000
J_Q04UKX2			5	Other British	00010000
J_Q04UKX2			6	Irish	00001000
J_Q04UKX2			7	Another white backgr	00000100
J_Q04UKX2			96	Valid skip	00000010
J_Q04UKX3	9	Background - Ethnic group - UK Welsh	-1	Missing	00000001
J_Q04UKX3			1	Welsh	00000000
J_Q04UKX3			2	English	10000000
J_Q04UKX3			3	Scottish	01000000
J_Q04UKX3			4	Northern Irish	00100000
J_Q04UKX3			5	Other British	00010000
J_Q04UKX3			6	Irish	00001000
J_Q04UKX3			7	Another white backgr	00000100
J_Q04UKX3			96	Valid skip	00000010
J_Q04UKX4	9	Background - Ethnic group - UK NI	-1	Missing	00000001
J_Q04UKX4			1	Northern Irish	00000000
J_Q04UKX4			2	English	10000000
J_Q04UKX4			3	Scottish	01000000
J_Q04UKX4			4	Welsh	00100000
J_Q04UKX4			5	Other British	00010000
J_Q04UKX4			6	Irish	00001000
J_Q04UKX4			7	Another white backgr	00000100
J_Q04UKX4			96	Valid skip	00000010
J_Q04UKX5	11	Background - Ethnic group - White origin	-1	Missing	0000000001
J_Q04UKX5			1	Scottish	0000000000
J_Q04UKX5			2	English	1000000000
J_Q04UKX5			3	Welsh	0100000000
J_Q04UKX5			4	Northern Irish	0010000000
J_Q04UKX5			5	British	0001000000
J_Q04UKX5			6	Irish	0000100000
J_Q04UKX5			7	Gypsy/Traveller	0000010000
J_Q04UKX5			8	Polish	0000001000
J_Q04UKX5			9	Another white backgr	0000000100
J_Q04UKX5			96	Valid skip	0000000010
J_Q04UKX6	6	Background - Ethnic group - White mixed ethnic	-1	Missing	00001
J_Q04UKX6			1	White+Black Caribbea	00000
J_Q04UKX6			2	White+Black African	10000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q04UKX6			3	White+Asian	01000
J_Q04UKX6			4	Another mixed backgr	00100
J_Q04UKX6			96	Valid skip	00010
J_Q04UKX7	7	Background - Ethnic group - Asian ethnic	-1	Missing	000001
J_Q04UKX7			1	Indian	000000
J_Q04UKX7			2	Pakistani	100000
J_Q04UKX7			3	Bangladeshi	010000
J_Q04UKX7			4	Chinese	001000
J_Q04UKX7			5	Other Asian Backgrou	000100
J_Q04UKX7			96	Valid skip	000010
J_Q04UKX8	5	Background - Ethnic group - Black ethnic	-1	Missing	0001
J_Q04UKX8			1	Caribbean	0000
J_Q04UKX8			2	African	1000
J_Q04UKX8			3	Another Black backgr	0100
J_Q04UKX8			96	Valid skip	0010
J_Q04UKX9	5	Background - Ethnic group - other ethnic (Eng, Wal	-1	Missing	0001
J_Q04UKX9			1	Arab	0000
J_Q04UKX9			2	Gypsy/Romany/Irish t	1000
J_Q04UKX9			3	Any other	0100
J_Q04UKX9			96	Valid skip	0010
J_Q05a1AT	25	Background - First learned language - NATIONAL	-1	Missing	000000000000000000000001
J_Q05a1AT			1	German	000000000000000000000000
J_Q05a1AT			2	Turkish	100000000000000000000000
J_Q05a1AT			3	Bosnian	010000000000000000000000
J_Q05a1AT			4	Croatian	001000000000000000000000
J_Q05a1AT			5	Serbian	000100000000000000000000
J_Q05a1AT			6	Arabic	000010000000000000000000
J_Q05a1AT			7	Chinese	000001000000000000000000
J_Q05a1AT			8	English	000000100000000000000000
J_Q05a1AT			9	French	000000010000000000000000
J_Q05a1AT			10	Italian	000000001000000000000000
J_Q05a1AT			11	Kurdish	000000000100000000000000
J_Q05a1AT			12	Macedonian	000000000010000000000000
J_Q05a1AT			13	Persian	000000000001000000000000
J_Q05a1AT			14	Polish	000000000000100000000000
J_Q05a1AT			15	Romanes	000000000000010000000000
J_Q05a1AT			16	Rumanian	000000000000001000000000
J_Q05a1AT			17	Slowakian	000000000000000100000000
J_Q05a1AT			18	Slovenian	000000000000000010000000
J_Q05a1AT			19	Spanish	000000000000000001000000
J_Q05a1AT			20	Swedish	000000000000000000100000
J_Q05a1AT			21	Czech	000000000000000000010000
J_Q05a1AT			22	Hungarian	000000000000000000001000
J_Q05a1AT			23	Other Lanugage	000000000000000000000100
J_Q05a1AT			96	Valid skip	000000000000000000000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a1AU	13	Background - First learned language (AUS)	-1	Missing	000000000001
J_Q05a1AU			1	English	000000000000
J_Q05a1AU			2	Italian	100000000000
J_Q05a1AU			3	Greek	010000000000
J_Q05a1AU			4	Cantonese	001000000000
J_Q05a1AU			5	Arabic	000100000000
J_Q05a1AU			6	Mandarin	000010000000
J_Q05a1AU			7	Vietnamese	000001000000
J_Q05a1AU			8	Spanish	000000100000
J_Q05a1AU			9	German	000000010000
J_Q05a1AU			10	Hindi	000000001000
J_Q05a1AU			11	Other	000000000100
J_Q05a1AU			96	Valid skip	000000000010
J_Q05a1AU6	7	Background - Reading skills in first language	-1	Missing	000001
J_Q05a1AU6			1	Excellent	000000
J_Q05a1AU6			2	Good	100000
J_Q05a1AU6			3	Moderate	010000
J_Q05a1AU6			4	Poor	001000
J_Q05a1AU6			5	Cannot read	000100
J_Q05a1AU6			6	Valid skip	000010
J_Q05a1AU7	7	Background - Writing skills in first language	-1	Missing	000001
J_Q05a1AU7			1	Excellent	000000
J_Q05a1AU7			2	Good	100000
J_Q05a1AU7			3	Moderate	010000
J_Q05a1AU7			4	Poor	001000
J_Q05a1AU7			5	Cannot write	000100
J_Q05a1AU7			6	Valid skip	000010
J_Q05a1AU8	7	Background - Reading skills in second language	-1	Missing	000001
J_Q05a1AU8			1	Excellent	000000
J_Q05a1AU8			2	Good	100000
J_Q05a1AU8			3	Moderate	010000
J_Q05a1AU8			4	Poor	001000
J_Q05a1AU8			5	Cannot read	000100
J_Q05a1AU8			6	Valid skip	000010
J_Q05a1AU9	7	Background - Writing skills in second language	-1	Missing	000001
J_Q05a1AU9			1	Excellent	000000
J_Q05a1AU9			2	Good	100000
J_Q05a1AU9			3	Moderate	010000
J_Q05a1AU9			4	Poor	001000
J_Q05a1AU9			5	Cannot write	000100
J_Q05a1AU9			6	Valid skip	000010
J_Q05a1AUa	4	Background - First learned language (AUS)	-1	Missing	001
J_Q05a1AUa			1	English	000
J_Q05a1AUa			2	Other	100
J_Q05a1AUa			96	Valid skip	010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a1BE	12	Background - First learned language	-1	Missing	00000000001
J_Q05a1BE			1	Dutch	00000000000
J_Q05a1BE			2	French	10000000000
J_Q05a1BE			3	German	01000000000
J_Q05a1BE			4	English	00100000000
J_Q05a1BE			5	Italian	00010000000
J_Q05a1BE			6	Spanish	00001000000
J_Q05a1BE			7	an Arabic language	00000100000
J_Q05a1BE			8	Turkish	00000010000
J_Q05a1BE			9	Polish	00000001000
J_Q05a1BE			10	Other	00000000100
J_Q05a1BE			96	Valid skip	00000000010
J_Q05a1CY	9	Background - First learned language	-1	Missing	00000001
J_Q05a1CY			1	Greek	00000000
J_Q05a1CY			2	English	10000000
J_Q05a1CY			3	Romanian	01000000
J_Q05a1CY			4	Russian	00100000
J_Q05a1CY			5	Armenian	00010000
J_Q05a1CY			6	Bulgarian	00001000
J_Q05a1CY			7	Other language	00000100
J_Q05a1CY			96	Valid skip	00000010
J_Q05a1CZ	9	Background - First learned language	-1	Missing	00000001
J_Q05a1CZ			1	Language1	00000000
J_Q05a1CZ			2	Language2	10000000
J_Q05a1CZ			3	Language3	01000000
J_Q05a1CZ			4	Language4	00100000
J_Q05a1CZ			5	Language5	00010000
J_Q05a1CZ			6	Language6	00001000
J_Q05a1CZ			7	Other language	00000100
J_Q05a1CZ			96	Valid skip	00000010
J_Q05a1DE	11	Background - First learned language	-1	Missing	0000000001
J_Q05a1DE			1	German	0000000000
J_Q05a1DE			2	Turkish	1000000000
J_Q05a1DE			3	Italian	0100000000
J_Q05a1DE			4	Polish	0010000000
J_Q05a1DE			5	Greek	0001000000
J_Q05a1DE			6	Serbian	0000100000
J_Q05a1DE			7	Croatian	0000010000
J_Q05a1DE			8	Russian	0000001000
J_Q05a1DE			9	Another language	0000000100
J_Q05a1DE			96	Valid skip	0000000010
J_Q05a1DK	10	Background - First learned language	-1	Missing	000000001
J_Q05a1DK			1	Danish	000000000
J_Q05a1DK			2	Turkish	100000000
J_Q05a1DK			3	German	010000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a1DK			4	Polish	001000000
J_Q05a1DK			5	Iraqi	000100000
J_Q05a1DK			6	Bosniaan	000010000
J_Q05a1DK			7	Norwegian	000001000
J_Q05a1DK			8	Other language	000000100
J_Q05a1DK			96	Valid skip	000000010
J_Q05a1EE	5	Background - First learned language	-1	Missing	0001
J_Q05a1EE			1	Estonian	0000
J_Q05a1EE			2	Russian	1000
J_Q05a1EE			3	Other, please specif	0100
J_Q05a1EE			96	Valid skip	0010
J_Q05a1ES	13	Background - First learned language	-1	Missing	000000000001
J_Q05a1ES			1	Not sn	000000000000
J_Q05a1ES			2	Nrabe	100000000000
J_Q05a1ES			3	Nrabeol	010000000000
J_Q05a1ES			4	Nrabeon	001000000000
J_Q05a1ES			5	Euskera	000100000000
J_Q05a1ES			6	Gallego	000010000000
J_Q05a1ES			7	Galles	000001000000
J_Q05a1ES			8	Quechuak	000000100000
J_Q05a1ES			9	Rumano	000000010000
J_Q05a1ES			10	Valenciano	000000001000
J_Q05a1ES			11	Otro idioma	000000000100
J_Q05a1ES			96	Valid skip	000000000010
J_Q05a1FI	11	Background - First learned language	-1	Missing	0000000001
J_Q05a1FI			1	Finnish	0000000000
J_Q05a1FI			2	Swedish	1000000000
J_Q05a1FI			3	Sami	0100000000
J_Q05a1FI			4	Romani	0010000000
J_Q05a1FI			5	Russian	0001000000
J_Q05a1FI			6	Estonian	0000100000
J_Q05a1FI			7	English	0000010000
J_Q05a1FI			8	German	0000001000
J_Q05a1FI			9	Other	0000000100
J_Q05a1FI			96	Valid skip	0000000010
J_Q05a1FR	12	Background - First learned language	-1	Missing	00000000001
J_Q05a1FR			1	French	00000000000
J_Q05a1FR			2	Regional language or	10000000000
J_Q05a1FR			3	Arabic	01000000000
J_Q05a1FR			4	German	00100000000
J_Q05a1FR			5	English	00010000000
J_Q05a1FR			6	Portuguese	00001000000
J_Q05a1FR			7	Italian	00000100000
J_Q05a1FR			8	Spanish	00000010000
J_Q05a1FR			9	Turkish	00000001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a1FR			10	Other. Please specif	00000000100
J_Q05a1FR			96	Valid skip	00000000010
J_Q05a1IE	10	Background - First learned language	-1	Missing	000000001
J_Q05a1IE			1	English	000000000
J_Q05a1IE			2	Irish	100000000
J_Q05a1IE			3	Polish	010000000
J_Q05a1IE			4	Lithuanian	001000000
J_Q05a1IE			5	Latvian	000100000
J_Q05a1IE			6	German	000010000
J_Q05a1IE			7	Romanian	000001000
J_Q05a1IE			8	Other	000000100
J_Q05a1IE			96	Valid skip	000000010
J_Q05a1IT	23	Background - First learned language	-1	Missing	0000000000000000000001
J_Q05a1IT			1	Italian	0000000000000000000000
J_Q05a1IT			2	Albanian	1000000000000000000000
J_Q05a1IT			3	Chinese	0100000000000000000000
J_Q05a1IT			4	English	0010000000000000000000
J_Q05a1IT			5	Filipino	0001000000000000000000
J_Q05a1IT			6	French	0000100000000000000000
J_Q05a1IT			7	German	0000010000000000000000
J_Q05a1IT			8	Moroccan	0000001000000000000000
J_Q05a1IT			9	Polish	0000000100000000000000
J_Q05a1IT			10	Romanian	0000000010000000000000
J_Q05a1IT			11	Romany (Gypsy)	0000000001000000000000
J_Q05a1IT			12	Spanish	0000000000100000000000
J_Q05a1IT			13	Tunisian Arabic	0000000000010000000000
J_Q05a1IT			14	Ukrainian	0000000000001000000000
J_Q05a1IT			15	Catalano	0000000000000100000000
J_Q05a1IT			16	Franco Provenzale	0000000000000010000000
J_Q05a1IT			17	Friulano	0000000000000001000000
J_Q05a1IT			18	Occitano	0000000000000000100000
J_Q05a1IT			19	Sardo	0000000000000000010000
J_Q05a1IT			20	Serbo-Croatian	0000000000000000001000
J_Q05a1IT			21	Other	0000000000000000000100
J_Q05a1IT			96	Valid skip	0000000000000000000010
J_Q05a1JP	9	Background - First learned language	-1	Missing	00000001
J_Q05a1JP			1	Japanese	00000000
J_Q05a1JP			2	Korean	10000000
J_Q05a1JP			3	Chinese	01000000
J_Q05a1JP			4	English	00100000
J_Q05a1JP			5	Portuguese	00010000
J_Q05a1JP			6	Spanish	00001000
J_Q05a1JP			7	Other language	00000100
J_Q05a1JP			96	Valid skip	00000010
J_Q05a1KO	9	KO_Background - First learned language	-1	Missing	00000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a1KO			1	Korean	00000000
J_Q05a1KO			2	Chinese	10000000
J_Q05a1KO			3	English	01000000
J_Q05a1KO			4	Vietnamese	00100000
J_Q05a1KO			5	Filipino	00010000
J_Q05a1KO			6	Japanese	00001000
J_Q05a1KO			7	Other language	00000100
J_Q05a1KO			96	Valid skip	00000010
J_Q05a1NL	9	Background - First learned language	-1	Missing	00000001
J_Q05a1NL			1	dutch	00000000
J_Q05a1NL			2	arabic	10000000
J_Q05a1NL			3	turkish	01000000
J_Q05a1NL			4	chinese	00100000
J_Q05a1NL			5	french	00010000
J_Q05a1NL			6	english	00001000
J_Q05a1NL			7	other language	00000100
J_Q05a1NL			96	Valid skip	00000010
J_Q05a1NO	21	Background - First learned language	-1	Missing	00000000000000000001
J_Q05a1NO			1	Norwegian	00000000000000000000
J_Q05a1NO			2	Danish	10000000000000000000
J_Q05a1NO			3	English	01000000000000000000
J_Q05a1NO			4	French	00100000000000000000
J_Q05a1NO			5	Hindi	00010000000000000000
J_Q05a1NO			6	Kurd	00001000000000000000
J_Q05a1NO			7	Persian	00000100000000000000
J_Q05a1NO			8	Punjabi	00000010000000000000
J_Q05a1NO			9	Serbian	00000001000000000000
J_Q05a1NO			10	Serbo-Croat	00000000100000000000
J_Q05a1NO			11	Singhalese	00000000010000000000
J_Q05a1NO			12	Somali	00000000001000000000
J_Q05a1NO			13	Spanish	00000000000100000000
J_Q05a1NO			14	Swedish	00000000000010000000
J_Q05a1NO			15	Turkish	00000000000001000000
J_Q05a1NO			16	German	00000000000000100000
J_Q05a1NO			17	Urdu	00000000000000010000
J_Q05a1NO			18	Vietnamese	00000000000000001000
J_Q05a1NO			19	Other language	00000000000000000100
J_Q05a1NO			96	Valid skip	00000000000000000010
J_Q05a1PL	14	Background - First learned language	-1	Missing	0000000000001
J_Q05a1PL			1	Byelorussian	0000000000000
J_Q05a1PL			2	Czech	1000000000000
J_Q05a1PL			3	Dutch	0100000000000
J_Q05a1PL			4	English	0010000000000
J_Q05a1PL			5	French	0001000000000
J_Q05a1PL			6	German	0000100000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a1PL			7	Lithuanian	0000010000000
J_Q05a1PL			8	Polish	0000001000000
J_Q05a1PL			9	Russian	0000000100000
J_Q05a1PL			10	Slovak	0000000010000
J_Q05a1PL			11	Ukrainian	0000000001000
J_Q05a1PL			12	Other language	0000000000100
J_Q05a1PL			96	Valid skip	0000000000010
J_Q05a1RU	9	Background - First learned language	-1	Missing	00000001
J_Q05a1RU			1	Language1	00000000
J_Q05a1RU			2	Language2	10000000
J_Q05a1RU			3	Language3	01000000
J_Q05a1RU			4	Language4	00100000
J_Q05a1RU			5	Language5	00010000
J_Q05a1RU			6	Language6	00001000
J_Q05a1RU			7	Other language	00000100
J_Q05a1RU			96	Valid skip	00000010
J_Q05a1SE	14	Background - First learned language	-1	Missing	0000000000001
J_Q05a1SE			1	Svenska	0000000000000
J_Q05a1SE			2	Finska	1000000000000
J_Q05a1SE			3	Spanska	0100000000000
J_Q05a1SE			4	Arabiska	0010000000000
J_Q05a1SE			5	Persiska	0001000000000
J_Q05a1SE			6	Polska	0000100000000
J_Q05a1SE			7	Serbokroatiska	0000010000000
J_Q05a1SE			8	Engelska	0000001000000
J_Q05a1SE			9	Turkiska	0000000100000
J_Q05a1SE			10	Bosniska	0000000010000
J_Q05a1SE			11	Kurdiska	0000000001000
J_Q05a1SE			12	Kurdiska k ange	0000000000100
J_Q05a1SE			96	Valid skip	0000000000010
J_Q05a1SK	9	Background - First learned language	-1	Missing	00000001
J_Q05a1SK			1	Slovak	00000000
J_Q05a1SK			2	Czech	10000000
J_Q05a1SK			3	Hungarian	01000000
J_Q05a1SK			4	German	00100000
J_Q05a1SK			5	Roma	00010000
J_Q05a1SK			6	Polish	00001000
J_Q05a1SK			7	Other language	00000100
J_Q05a1SK			96	Valid skip	00000010
J_Q05a1UK	12	Background - First learned language	-1	Missing	00000000001
J_Q05a1UK			1	English	00000000000
J_Q05a1UK			2	Welsh	10000000000
J_Q05a1UK			3	Irish	01000000000
J_Q05a1UK			4	Scottish Gaelic	00100000000
J_Q05a1UK			5	Ulster Scots/Ullans	00010000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a1UK			6	Hindi	00001000000
J_Q05a1UK			7	Urdu	00000100000
J_Q05a1UK			8	Punjabi	00000010000
J_Q05a1UK			9	Polish	00000001000
J_Q05a1UK			10	Other	00000000100
J_Q05a1UK			96	Valid skip	00000000010
J_Q05a1US	9	Background - First learned language	-1	Missing	00000001
J_Q05a1US			1	English	00000000
J_Q05a1US			2	Spanis	10000000
J_Q05a1US			3	French	01000000
J_Q05a1US			4	Italian	00100000
J_Q05a1US			5	Chinese	00010000
J_Q05a1US			6	German	00001000
J_Q05a1US			7	Other language	00000100
J_Q05a1US			96	Valid skip	00000010
J_Q05a2AT	25	Background - Second learned language - NATIONAL	-1	Missing	000000000000000000000001
J_Q05a2AT			1	German	000000000000000000000000
J_Q05a2AT			2	Turkish	100000000000000000000000
J_Q05a2AT			3	Bosnian	010000000000000000000000
J_Q05a2AT			4	Croatian	001000000000000000000000
J_Q05a2AT			5	Serbian	000100000000000000000000
J_Q05a2AT			6	Arabic	000010000000000000000000
J_Q05a2AT			7	Chinese	000001000000000000000000
J_Q05a2AT			8	English	000000100000000000000000
J_Q05a2AT			9	French	000000010000000000000000
J_Q05a2AT			10	Italian	000000001000000000000000
J_Q05a2AT			11	Kurdish	000000000100000000000000
J_Q05a2AT			12	Macedonian	000000000010000000000000
J_Q05a2AT			13	Persian	000000000001000000000000
J_Q05a2AT			14	Polish	000000000000100000000000
J_Q05a2AT			15	Romanes	000000000000010000000000
J_Q05a2AT			16	Rumanian	000000000000001000000000
J_Q05a2AT			17	Slowakian	000000000000000100000000
J_Q05a2AT			18	Slovenian	000000000000000010000000
J_Q05a2AT			19	Spanish	000000000000000001000000
J_Q05a2AT			20	Swedish	000000000000000000100000
J_Q05a2AT			21	Czech	000000000000000000010000
J_Q05a2AT			22	Hungarian	000000000000000000001000
J_Q05a2AT			23	Other Lanugage	000000000000000000000100
J_Q05a2AT			96	Valid skip	000000000000000000000010
J_Q05a2ATX	6	Background - Second learned language - skill - NAT	-1	Missing	00001
J_Q05a2ATX			1	I just understand fe	00000
J_Q05a2ATX			2	I can use the most c	10000
J_Q05a2ATX			3	I understand the mai	01000
J_Q05a2ATX			4	I can use the langua	00100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a2ATX			6	Valid skip	00010
J_Q05a2AU	13	Background - Second learned language (AUS)	-1	Missing	000000000001
J_Q05a2AU			1	English	000000000000
J_Q05a2AU			2	Italian	100000000000
J_Q05a2AU			3	Greek	010000000000
J_Q05a2AU			4	Cantonese	001000000000
J_Q05a2AU			5	Arabic	000100000000
J_Q05a2AU			6	Mandarin	000010000000
J_Q05a2AU			7	Vietnamese	000001000000
J_Q05a2AU			8	Spanish	000000100000
J_Q05a2AU			9	German	000000010000
J_Q05a2AU			10	Hindi	000000001000
J_Q05a2AU			11	Other	000000000100
J_Q05a2AU			96	Valid skip	000000000010
J_Q05a2AUa	4	Background - Second learned language (AUS)	-1	Missing	001
J_Q05a2AUa			1	English	000
J_Q05a2AUa			2	Other	100
J_Q05a2AUa			96	Valid skip	010
J_Q05a2BE	12	Background - Second learned language	-1	Missing	00000000001
J_Q05a2BE			1	Dutch	00000000000
J_Q05a2BE			2	French	10000000000
J_Q05a2BE			3	German	01000000000
J_Q05a2BE			4	English	00100000000
J_Q05a2BE			5	Italian	00010000000
J_Q05a2BE			6	Spanish	00001000000
J_Q05a2BE			7	an Arabic language	00000100000
J_Q05a2BE			8	Turkish	00000010000
J_Q05a2BE			9	Polish	00000001000
J_Q05a2BE			10	Other	00000000100
J_Q05a2BE			96	Valid skip	00000000010
J_Q05a2CY	9	Background - Second learned language	-1	Missing	00000001
J_Q05a2CY			1	Greek	00000000
J_Q05a2CY			2	English	10000000
J_Q05a2CY			3	Turkish	01000000
J_Q05a2CY			4	Russian	00100000
J_Q05a2CY			5	Armenian	00010000
J_Q05a2CY			6	Bulgarian	00001000
J_Q05a2CY			7	Other language	00000100
J_Q05a2CY			96	Valid skip	00000010
J_Q05a2CZ	9	Background - Second learned language	-1	Missing	00000001
J_Q05a2CZ			1	Language1	00000000
J_Q05a2CZ			2	Language2	10000000
J_Q05a2CZ			3	Language3	01000000
J_Q05a2CZ			4	Language4	00100000
J_Q05a2CZ			5	Language5	00010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a2CZ			6	Language6	00001000
J_Q05a2CZ			7	Other language	00000100
J_Q05a2CZ			96	Valid skip	00000010
J_Q05a2DE	11	Background - Second learned language	-1	Missing	0000000001
J_Q05a2DE			1	German	0000000000
J_Q05a2DE			2	Turkish	1000000000
J_Q05a2DE			3	Italian	0100000000
J_Q05a2DE			4	Polish	0010000000
J_Q05a2DE			5	Greek	0001000000
J_Q05a2DE			6	Serbian	0000100000
J_Q05a2DE			7	Croatian	0000010000
J_Q05a2DE			8	Russian	0000001000
J_Q05a2DE			9	Another language	0000000100
J_Q05a2DE			96	Valid skip	0000000010
J_Q05a2DK	10	Background - Second learned language	-1	Missing	000000001
J_Q05a2DK			1	Danish	000000000
J_Q05a2DK			2	Turkish	100000000
J_Q05a2DK			3	German	010000000
J_Q05a2DK			4	Polish	001000000
J_Q05a2DK			5	Iraqi	000100000
J_Q05a2DK			6	Bosniaan	000010000
J_Q05a2DK			7	Norwegian	000001000
J_Q05a2DK			8	Other language	000000100
J_Q05a2DK			96	Valid skip	000000010
J_Q05a2EE	5	Background - Second learned language	-1	Missing	0001
J_Q05a2EE			1	Estonian	0000
J_Q05a2EE			2	Russian	1000
J_Q05a2EE			3	Other, please specif	0100
J_Q05a2EE			96	Valid skip	0010
J_Q05a2ES	13	Background - Second learned language	-1	Missing	000000000001
J_Q05a2ES			1	Not sn	000000000000
J_Q05a2ES			2	Nrabe	100000000000
J_Q05a2ES			3	Nrabeol	010000000000
J_Q05a2ES			4	Nrabeon	001000000000
J_Q05a2ES			5	Euskera	000100000000
J_Q05a2ES			6	Gallego	000010000000
J_Q05a2ES			7	Ingles	000001000000
J_Q05a2ES			8	Quechuak	000000100000
J_Q05a2ES			9	Rumano	000000010000
J_Q05a2ES			10	Valenciano	000000001000
J_Q05a2ES			11	Otro idioma	000000000100
J_Q05a2ES			96	Valid skip	000000000010
J_Q05a2FI	11	Background - Second learned language	-1	Missing	0000000001
J_Q05a2FI			1	Finnish	0000000000
J_Q05a2FI			2	Swedish	1000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a2FI			3	Sami	0100000000
J_Q05a2FI			4	Romani	0010000000
J_Q05a2FI			5	Russian	0001000000
J_Q05a2FI			6	Estonian	0000100000
J_Q05a2FI			7	English	0000010000
J_Q05a2FI			8	German	0000001000
J_Q05a2FI			9	Other	0000000100
J_Q05a2FI			96	Valid skip	0000000010
J_Q05a2FR	12	Background - Second learned language	-1	Missing	00000000001
J_Q05a2FR			1	French	00000000000
J_Q05a2FR			2	Regional language or	10000000000
J_Q05a2FR			3	Arabic	01000000000
J_Q05a2FR			4	German	00100000000
J_Q05a2FR			5	English	00010000000
J_Q05a2FR			6	Portuguese	00001000000
J_Q05a2FR			7	Italian	00000100000
J_Q05a2FR			8	Spanish	00000010000
J_Q05a2FR			9	Turkish	00000001000
J_Q05a2FR			10	Other. Please specif	00000000100
J_Q05a2FR			96	Valid skip	00000000010
J_Q05a2IE	10	Background - Second learned language	-1	Missing	000000001
J_Q05a2IE			1	English	000000000
J_Q05a2IE			2	Irish	100000000
J_Q05a2IE			3	Polish	010000000
J_Q05a2IE			4	Lithuanian	001000000
J_Q05a2IE			5	Latvian	000100000
J_Q05a2IE			6	German	000010000
J_Q05a2IE			7	Romanian	000001000
J_Q05a2IE			8	Other	000000100
J_Q05a2IE			96	Valid skip	000000010
J_Q05a2IT	23	Background - Second learned language	-1	Missing	0000000000000000000001
J_Q05a2IT			1	Italian	0000000000000000000000
J_Q05a2IT			2	Albanian	1000000000000000000000
J_Q05a2IT			3	Chinese	0100000000000000000000
J_Q05a2IT			4	English	0010000000000000000000
J_Q05a2IT			5	Filipino	0001000000000000000000
J_Q05a2IT			6	French	0000100000000000000000
J_Q05a2IT			7	German	0000010000000000000000
J_Q05a2IT			8	Moroccan	0000001000000000000000
J_Q05a2IT			9	Polish	0000000100000000000000
J_Q05a2IT			10	Romanian	0000000010000000000000
J_Q05a2IT			11	Romany (Gypsy)	0000000001000000000000
J_Q05a2IT			12	Spanish	0000000000100000000000
J_Q05a2IT			13	Tunisian Arabic	0000000000010000000000
J_Q05a2IT			14	Ukrainian	0000000000001000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a2IT			15	Catalano	0000000000000100000000
J_Q05a2IT			16	Franco Provenzale	0000000000000010000000
J_Q05a2IT			17	Friulano	0000000000000001000000
J_Q05a2IT			18	Occitano	0000000000000000100000
J_Q05a2IT			19	Sardo	0000000000000000010000
J_Q05a2IT			20	Serbo-Croatian	0000000000000000001000
J_Q05a2IT			21	Other	0000000000000000000100
J_Q05a2IT			96	Valid skip	0000000000000000000010
J_Q05a2JP	9	Background - Second learned language	-1	Missing	00000001
J_Q05a2JP			1	Japanese	00000000
J_Q05a2JP			2	Korean	10000000
J_Q05a2JP			3	Chinese	01000000
J_Q05a2JP			4	English	00100000
J_Q05a2JP			5	Portuguese	00010000
J_Q05a2JP			6	Spanish	00001000
J_Q05a2JP			7	Other language	00000100
J_Q05a2JP			96	Valid skip	00000010
J_Q05a2KO	9	KO_Background - Second learned language	-1	Missing	00000001
J_Q05a2KO			1	Korean	00000000
J_Q05a2KO			2	Chinese	10000000
J_Q05a2KO			3	English	01000000
J_Q05a2KO			4	Vietnamese	00100000
J_Q05a2KO			5	Filipino	00010000
J_Q05a2KO			6	Japanese	00001000
J_Q05a2KO			7	Other language	00000100
J_Q05a2KO			96	Valid skip	00000010
J_Q05a2NL	9	Background - Second learned language	-1	Missing	00000001
J_Q05a2NL			1	dutch	00000000
J_Q05a2NL			2	arabic	10000000
J_Q05a2NL			3	turkish	01000000
J_Q05a2NL			4	chinese	00100000
J_Q05a2NL			5	french	00010000
J_Q05a2NL			6	english	00001000
J_Q05a2NL			7	other language	00000100
J_Q05a2NL			96	Valid skip	00000010
J_Q05a2NO	21	Background - Second learned language	-1	Missing	00000000000000000001
J_Q05a2NO			1	Norwegian	00000000000000000000
J_Q05a2NO			2	Danish	10000000000000000000
J_Q05a2NO			3	English	01000000000000000000
J_Q05a2NO			4	French	00100000000000000000
J_Q05a2NO			5	Hindi	00010000000000000000
J_Q05a2NO			6	Kurd	00001000000000000000
J_Q05a2NO			7	Persian	00000100000000000000
J_Q05a2NO			8	Punjabi	00000010000000000000
J_Q05a2NO			9	Serbian	00000001000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a2NO			10	Serbo-Croat	00000000100000000000
J_Q05a2NO			11	Singhalese	00000000010000000000
J_Q05a2NO			12	Somali	00000000001000000000
J_Q05a2NO			13	Spanish	00000000000100000000
J_Q05a2NO			14	Swedish	00000000000010000000
J_Q05a2NO			15	Turkish	00000000000001000000
J_Q05a2NO			16	German	00000000000000100000
J_Q05a2NO			17	Urdu	00000000000000010000
J_Q05a2NO			18	Vietnamese	00000000000000001000
J_Q05a2NO			19	Other language	00000000000000000100
J_Q05a2NO			186	Valid skip	00000000000000000010
J_Q05a2PL	14	Background - Second learned language	-1	Missing	0000000000001
J_Q05a2PL			1	Byelorussian	0000000000000
J_Q05a2PL			2	Czech	1000000000000
J_Q05a2PL			3	Dutch	0100000000000
J_Q05a2PL			4	English	0010000000000
J_Q05a2PL			5	French	0001000000000
J_Q05a2PL			6	German	0000100000000
J_Q05a2PL			7	Lithuanian	0000010000000
J_Q05a2PL			8	Polish	0000001000000
J_Q05a2PL			9	Russian	0000000100000
J_Q05a2PL			10	Slovak	0000000010000
J_Q05a2PL			11	Ukrainian	0000000001000
J_Q05a2PL			12	Other language	0000000000100
J_Q05a2PL			96	Valid skip	0000000000010
J_Q05a2RU	9	Background - Second learned language	-1	Missing	00000001
J_Q05a2RU			1	Language1	00000000
J_Q05a2RU			2	Language2	10000000
J_Q05a2RU			3	Language3	01000000
J_Q05a2RU			4	Language4	00100000
J_Q05a2RU			5	Language5	00010000
J_Q05a2RU			6	Language6	00001000
J_Q05a2RU			7	Other language	00000100
J_Q05a2RU			96	Valid skip	00000010
J_Q05a2SE	14	Background - Second learned language	-1	Missing	0000000000001
J_Q05a2SE			1	Svenska	0000000000000
J_Q05a2SE			2	Finska	1000000000000
J_Q05a2SE			3	Spanska	0100000000000
J_Q05a2SE			4	Arabiska	0010000000000
J_Q05a2SE			5	Persiska	0001000000000
J_Q05a2SE			6	Polska	0000100000000
J_Q05a2SE			7	Serbokroatiska	0000010000000
J_Q05a2SE			8	Engelska	0000001000000
J_Q05a2SE			9	Turkiska	0000000100000
J_Q05a2SE			10	Bosniska	0000000010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05a2SE			11	Kurdiska	0000000001000
J_Q05a2SE			12	Kurdiska k ange	0000000000100
J_Q05a2SE			96	Valid skip	0000000000010
J_Q05a2SK	9	Background - Second learned language	-1	Missing	00000001
J_Q05a2SK			1	Slovak	00000000
J_Q05a2SK			2	Czech	10000000
J_Q05a2SK			3	Hungarian	01000000
J_Q05a2SK			4	German	00100000
J_Q05a2SK			5	Roma	00010000
J_Q05a2SK			6	Polish	00001000
J_Q05a2SK			7	Other language	00000100
J_Q05a2SK			96	Valid skip	00000010
J_Q05a2UK	12	Background - Second learned language	-1	Missing	00000000001
J_Q05a2UK			1	English	00000000000
J_Q05a2UK			2	Welsh	10000000000
J_Q05a2UK			3	Irish	01000000000
J_Q05a2UK			4	Scottish Gaelic	00100000000
J_Q05a2UK			5	Ulster Scots/Ullans	00010000000
J_Q05a2UK			6	Hindi	00001000000
J_Q05a2UK			7	Urdu	00000100000
J_Q05a2UK			8	Punjabi	00000010000
J_Q05a2UK			9	Polish	00000001000
J_Q05a2UK			10	Other	00000000100
J_Q05a2UK			96	Valid skip	00000000010
J_Q05a2US	9	Background - Second learned language	-1	Missing	00000001
J_Q05a2US			1	English	00000000
J_Q05a2US			2	Spanis	10000000
J_Q05a2US			3	French	01000000
J_Q05a2US			4	Italian	00100000
J_Q05a2US			5	Chinese	00010000
J_Q05a2US			6	German	00001000
J_Q05a2US			7	Other language	00000100
J_Q05a2US			96	Valid skip	00000010
J_Q05a2USX2	8	Background - Age learned English	-1	Missing	0000001
J_Q05a2USX2			1	1-4 years old	0000000
J_Q05a2USX2			2	5-10 years old	1000000
J_Q05a2USX2			3	11-15 years old	0100000
J_Q05a2USX2			4	16-20 years old	0010000
J_Q05a2USX2			5	21 years or older	0001000
J_Q05a2USX2			6	Does not speak Engli	0000100
J_Q05a2USX2			96	Valid skip	0000010
J_Q05bAT	25	Background - Language spoken at home - NATIONAL	-1	Missing	000000000000000000000001
J_Q05bAT			1	German	000000000000000000000000
J_Q05bAT			2	Turkish	100000000000000000000000
J_Q05bAT			3	Bosnian	010000000000000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05bAT			4	Croatian	001000000000000000000000
J_Q05bAT			5	Serbian	000100000000000000000000
J_Q05bAT			6	Arabic	000010000000000000000000
J_Q05bAT			7	Chinese	000001000000000000000000
J_Q05bAT			8	English	000000100000000000000000
J_Q05bAT			9	French	000000010000000000000000
J_Q05bAT			10	Italian	000000001000000000000000
J_Q05bAT			11	Kurdish	000000000100000000000000
J_Q05bAT			12	Macedonian	000000000010000000000000
J_Q05bAT			13	Persian	000000000001000000000000
J_Q05bAT			14	Polish	000000000000100000000000
J_Q05bAT			15	Romanes	000000000000010000000000
J_Q05bAT			16	Rumanian	000000000000001000000000
J_Q05bAT			17	Slowakian	000000000000000100000000
J_Q05bAT			18	Slovenian	000000000000000010000000
J_Q05bAT			19	Spanish	000000000000000001000000
J_Q05bAT			20	Swedish	000000000000000000100000
J_Q05bAT			21	Czech	000000000000000000010000
J_Q05bAT			22	Hungarian	000000000000000000001000
J_Q05bAT			23	Other Lanugage	000000000000000000000100
J_Q05bAT			96	Valid skip	000000000000000000000010
J_Q05bATX1	26	Background - Language beside mother tongue - NATIO	-1	Missing	0000000000000000000000001
J_Q05bATX1			0	No further language	0000000000000000000000000
J_Q05bATX1			1	German	1000000000000000000000000
J_Q05bATX1			2	Turkish	0100000000000000000000000
J_Q05bATX1			3	Bosnian	0010000000000000000000000
J_Q05bATX1			4	Croatian	0001000000000000000000000
J_Q05bATX1			5	Serbian	0000100000000000000000000
J_Q05bATX1			6	Arabic	0000010000000000000000000
J_Q05bATX1			7	Chinese	0000001000000000000000000
J_Q05bATX1			8	English	0000000100000000000000000
J_Q05bATX1			9	French	0000000010000000000000000
J_Q05bATX1			10	Italian	0000000001000000000000000
J_Q05bATX1			11	Kurdish	0000000000100000000000000
J_Q05bATX1			12	Macedonian	0000000000010000000000000
J_Q05bATX1			13	Persian	0000000000001000000000000
J_Q05bATX1			14	Polish	0000000000000100000000000
J_Q05bATX1			15	Romanes	0000000000000010000000000
J_Q05bATX1			16	Rumanian	0000000000000001000000000
J_Q05bATX1			17	Slowakian	0000000000000000100000000
J_Q05bATX1			18	Slovenian	0000000000000000010000000
J_Q05bATX1			19	Spanish	0000000000000000001000000
J_Q05bATX1			20	Swedish	0000000000000000000100000
J_Q05bATX1			21	Czech	0000000000000000000010000
J_Q05bATX1			22	Hungarian	0000000000000000000001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05bATX1			23	Other Lanugage	000000000000000000000100
J_Q05bATX1			96	Valid skip	000000000000000000000010
J_Q05bATX2	6	Background - Language beside mother tongue - skill	-1	Missing	00001
J_Q05bATX2			1	I just understand fe	00000
J_Q05bATX2			2	I can use the most c	10000
J_Q05bATX2			3	I understand the mai	01000
J_Q05bATX2			4	I can use the langua	00100
J_Q05bATX2			6	Valid skip	00010
J_Q05bAU1	13	Background - Language mainly spoken at home (AUS)	-1	Missing	000000000001
J_Q05bAU1			1	English	000000000000
J_Q05bAU1			2	Italian	100000000000
J_Q05bAU1			3	Greek	010000000000
J_Q05bAU1			4	Cantonese	001000000000
J_Q05bAU1			5	Arabic	000100000000
J_Q05bAU1			6	Mandarin	000010000000
J_Q05bAU1			7	Vietnamese	000001000000
J_Q05bAU1			8	Spanish	000000100000
J_Q05bAU1			9	German	000000010000
J_Q05bAU1			10	Hindi	000000001000
J_Q05bAU1			11	Other	000000000100
J_Q05bAU1			96	Valid skip	000000000010
J_Q05bAU1a	4	Background - Language mainly spoken at home (AUS)	-1	Missing	001
J_Q05bAU1a			1	English	000
J_Q05bAU1a			2	Other	100
J_Q05bAU1a			96	Valid skip	010
J_Q05bAU3	6	Background - Rate speaking english	-1	Missing	00001
J_Q05bAU3			1	Very well	00000
J_Q05bAU3			2	Well	10000
J_Q05bAU3			3	Not well	01000
J_Q05bAU3			4	Not at all	00100
J_Q05bAU3			6	Valid skip	00010
J_Q05bBE	12	Background - Language spoken at home	-1	Missing	00000000001
J_Q05bBE			1	Dutch	00000000000
J_Q05bBE			2	French	10000000000
J_Q05bBE			3	German	01000000000
J_Q05bBE			4	English	00100000000
J_Q05bBE			5	Italian	00010000000
J_Q05bBE			6	Spanish	00001000000
J_Q05bBE			7	an Arabic language	00000100000
J_Q05bBE			8	Turkish	00000010000
J_Q05bBE			9	Polish	00000001000
J_Q05bBE			10	Other	00000000100
J_Q05bBE			96	Valid skip	00000000010
J_Q05bCY	9	Background - Language spoken at home	-1	Missing	00000001
J_Q05bCY			1	Greek	00000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05bCY			2	English	10000000
J_Q05bCY			3	Turkish	01000000
J_Q05bCY			4	Russian	00100000
J_Q05bCY			5	Armenian	00010000
J_Q05bCY			6	Bulgarian	00001000
J_Q05bCY			7	Other language	00000100
J_Q05bCY			96	Valid skip	00000010
J_Q05bCZ	9	Background - Language spoken at home	-1	Missing	00000001
J_Q05bCZ			1	Language1	00000000
J_Q05bCZ			2	Language2	10000000
J_Q05bCZ			3	Language3	01000000
J_Q05bCZ			4	Language4	00100000
J_Q05bCZ			5	Language5	00010000
J_Q05bCZ			6	Language6	00001000
J_Q05bCZ			7	Other language	00000100
J_Q05bCZ			96	Valid skip	00000010
J_Q05bDE	11	Background - Language spoken at home	-1	Missing	0000000001
J_Q05bDE			1	German	0000000000
J_Q05bDE			2	Turkish	1000000000
J_Q05bDE			3	Italian	0100000000
J_Q05bDE			4	Polish	0010000000
J_Q05bDE			5	Greek	0001000000
J_Q05bDE			6	Serbian	0000100000
J_Q05bDE			7	Croatian	0000010000
J_Q05bDE			8	Russian	0000001000
J_Q05bDE			9	Another language	0000000100
J_Q05bDE			96	Valid skip	0000000010
J_Q05bDEX1	11	Background - Second language spoken at home	-1	Missing	0000000001
J_Q05bDEX1			1	German	0000000000
J_Q05bDEX1			2	Turkish	1000000000
J_Q05bDEX1			3	Italian	0100000000
J_Q05bDEX1			4	Polish	0010000000
J_Q05bDEX1			5	Greek	0001000000
J_Q05bDEX1			6	Serbian	0000100000
J_Q05bDEX1			7	Croatian	0000010000
J_Q05bDEX1			8	Russian	0000001000
J_Q05bDEX1			9	Another language	0000000100
J_Q05bDEX1			96	Valid skip	0000000010
J_Q05bDEX2	11	Background - Language at age 16	-1	Missing	0000000001
J_Q05bDEX2			1	German	0000000000
J_Q05bDEX2			2	Turkish	1000000000
J_Q05bDEX2			3	Italian	0100000000
J_Q05bDEX2			4	Polish	0010000000
J_Q05bDEX2			5	Greek	0001000000
J_Q05bDEX2			6	Serbian	0000100000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05bDEX2			7	Croatian	0000010000
J_Q05bDEX2			8	Russian	0000001000
J_Q05bDEX2			9	Another language	0000000100
J_Q05bDEX2			96	Valid skip	0000000010
J_Q05bDEX3	11	Background - Second language spoken at age 16	-1	Missing	0000000001
J_Q05bDEX3			1	German	0000000000
J_Q05bDEX3			2	Turkish	1000000000
J_Q05bDEX3			3	Italian	0100000000
J_Q05bDEX3			4	Polish	0010000000
J_Q05bDEX3			5	Greek	0001000000
J_Q05bDEX3			6	Serbian	0000100000
J_Q05bDEX3			7	Croatian	0000010000
J_Q05bDEX3			8	Russian	0000001000
J_Q05bDEX3			9	Another language	0000000100
J_Q05bDEX3			96	Valid skip	0000000010
J_Q05bDK	10	Background - Language spoken at home	-1	Missing	000000001
J_Q05bDK			1	Danish	000000000
J_Q05bDK			2	Turkish	100000000
J_Q05bDK			3	German	010000000
J_Q05bDK			4	Polish	001000000
J_Q05bDK			5	Iraqi	000100000
J_Q05bDK			6	Bosniaan	000010000
J_Q05bDK			7	Norwegian	000001000
J_Q05bDK			8	Other language	000000100
J_Q05bDK			96	Valid skip	000000010
J_Q05bDKx1	7	last 12 months, how often have you used the langua	-1	Missing	000001
J_Q05bDKx1			1	Never	000000
J_Q05bDKx1			2	Less than once a mon	100000
J_Q05bDKx1			3	Less than once a wee	010000
J_Q05bDKx1			4	At least once a week	001000
J_Q05bDKx1			5	Every day	000100
J_Q05bDKx1			96	Valid skip	000010
J_Q05bDKx2	7	last 12 months, how often have you used other fore	-1	Missing	000001
J_Q05bDKx2			1	Never	000000
J_Q05bDKx2			2	Less than once a mon	100000
J_Q05bDKx2			3	Less than once a wee	010000
J_Q05bDKx2			4	At least once a week	001000
J_Q05bDKx2			5	Every day	000100
J_Q05bDKx2			96	Valid skip	000010
J_Q05bEE	5	Background - Language spoken at home	-1	Missing	0001
J_Q05bEE			1	Estonian	0000
J_Q05bEE			2	Russian	1000
J_Q05bEE			3	Other, please specif	0100
J_Q05bEE			96	Valid skip	0010
J_Q05bEEX1	7	Background - Proficiency in this (Estonian/Russian	-1	Missing	000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05bEEX1			1	Not at all	000000
J_Q05bEEX1			2	I command on the ver	100000
J_Q05bEEX1			3	I manage with some p	010000
J_Q05bEEX1			4	I manage well on the	001000
J_Q05bEEX1			5	I am fluent in the l	000100
J_Q05bEEX1			6	Valid skip	000010
J_Q05bEEX2	5	Background - Any other language	-1	Missing	0001
J_Q05bEEX2			1	Yes, one other langu	0000
J_Q05bEEX2			2	Yes, more than one o	1000
J_Q05bEEX2			3	No	0100
J_Q05bEEX2			96	Valid skip	0010
J_Q05bEEX4	6	Background - Proficiency in this language	-1	Missing	00001
J_Q05bEEX4			1	I command on the ver	00000
J_Q05bEEX4			2	I manage with some p	10000
J_Q05bEEX4			3	I manage well on the	01000
J_Q05bEEX4			4	I am fluent in the l	00100
J_Q05bEEX4			6	Valid skip	00010
J_Q05bEEX6	6	Background - Proficiency in this language	-1	Missing	00001
J_Q05bEEX6			1	I command on the ver	00000
J_Q05bEEX6			2	I manage with some p	10000
J_Q05bEEX6			3	I manage well on the	01000
J_Q05bEEX6			4	I am fluent in the l	00100
J_Q05bEEX6			6	Valid skip	00010
J_Q05bEEX8	6	Background - Proficiency in this language	-1	Missing	00001
J_Q05bEEX8			1	I command on the ver	00000
J_Q05bEEX8			2	I manage with some p	10000
J_Q05bEEX8			3	I manage well on the	01000
J_Q05bEEX8			4	I am fluent in the l	00100
J_Q05bEEX8			6	Valid skip	00010
J_Q05bES	13	Background - Language spoken at home	-1	Missing	000000000001
J_Q05bES			1	Not sn	000000000000
J_Q05bES			2	Nrabe	100000000000
J_Q05bES			3	Nrabeol	010000000000
J_Q05bES			4	Nrabeon	001000000000
J_Q05bES			5	Euskera	000100000000
J_Q05bES			6	Gallego	000010000000
J_Q05bES			7	Galles	000001000000
J_Q05bES			8	Quechuak	000000100000
J_Q05bES			9	Rumano	000000010000
J_Q05bES			10	Valenciano	000000001000
J_Q05bES			11	Other language	000000000100
J_Q05bES			96	Valid skip	000000000010
J_Q05bFI	11	Background - Language spoken at home	-1	Missing	0000000001
J_Q05bFI			1	Finnish	0000000000
J_Q05bFI			2	Swedish	1000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05bFI			3	Sami	0100000000
J_Q05bFI			4	Romani	0010000000
J_Q05bFI			5	Russian	0001000000
J_Q05bFI			6	Estonian	0000100000
J_Q05bFI			7	English	0000010000
J_Q05bFI			8	German	0000001000
J_Q05bFI			9	Other	0000000100
J_Q05bFI			96	Valid skip	0000000010
J_Q05bFR	12	Background - Language spoken at home	-1	Missing	00000000001
J_Q05bFR			1	French	00000000000
J_Q05bFR			2	Regional language or	10000000000
J_Q05bFR			3	Arabic	01000000000
J_Q05bFR			4	German	00100000000
J_Q05bFR			5	English	00010000000
J_Q05bFR			6	Portuguese	00001000000
J_Q05bFR			7	Italian	00000100000
J_Q05bFR			8	Spanish	00000010000
J_Q05bFR			9	Turkish	00000001000
J_Q05bFR			10	Other. Please specif	00000000100
J_Q05bFR			96	Valid skip	00000000010
J_Q05bIE	10	Background - Language spoken at home	-1	Missing	000000001
J_Q05bIE			1	English	000000000
J_Q05bIE			2	Irish	100000000
J_Q05bIE			3	Polish	010000000
J_Q05bIE			4	Lithuanian	001000000
J_Q05bIE			5	Latvian	000100000
J_Q05bIE			6	German	000010000
J_Q05bIE			7	Romanian	000001000
J_Q05bIE			8	Other	000000100
J_Q05bIE			96	Valid skip	000000010
J_Q05bIT	23	Background - Language spoken at home	-1	Missing	0000000000000000000001
J_Q05bIT			1	Italian	0000000000000000000000
J_Q05bIT			2	Albanian	1000000000000000000000
J_Q05bIT			3	Chinese	0100000000000000000000
J_Q05bIT			4	English	0010000000000000000000
J_Q05bIT			5	Filipino	0001000000000000000000
J_Q05bIT			6	French	0000100000000000000000
J_Q05bIT			7	German	0000010000000000000000
J_Q05bIT			8	Moroccan	0000001000000000000000
J_Q05bIT			9	Polish	0000000100000000000000
J_Q05bIT			10	Romanian	0000000010000000000000
J_Q05bIT			11	Romany (Gypsy)	0000000001000000000000
J_Q05bIT			12	Spanish	0000000000100000000000
J_Q05bIT			13	Tunisian Arabic	0000000000010000000000
J_Q05bIT			14	Ukrainian	0000000000001000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05bIT			15	Catalano	0000000000000100000000
J_Q05bIT			16	Franco Provenzale	0000000000000010000000
J_Q05bIT			17	Friulano	0000000000000001000000
J_Q05bIT			18	Occitano	0000000000000000100000
J_Q05bIT			19	Sardo	0000000000000000010000
J_Q05bIT			20	Serbo-Croatian	0000000000000000001000
J_Q05bIT			21	Other	0000000000000000000100
J_Q05bIT			96	Valid skip	0000000000000000000010
J_Q05bJP	9	Background - Language spoken at home	-1	Missing	00000001
J_Q05bJP			1	Japanese	00000000
J_Q05bJP			2	Korean	10000000
J_Q05bJP			3	Chinese	01000000
J_Q05bJP			4	English	00100000
J_Q05bJP			5	Portuguese	00010000
J_Q05bJP			6	Spanish	00001000
J_Q05bJP			7	Other language	00000100
J_Q05bJP			96	Valid skip	00000010
J_Q05bJPX	8	Background - Experience of living abroad	-1	Missing	0000001
J_Q05bJPX			1	Never	0000000
J_Q05bJPX			2	Less than 1 year	1000000
J_Q05bJPX			3	1 to 2 years	0100000
J_Q05bJPX			4	2 to 5 years	0010000
J_Q05bJPX			5	5 to 10 years	0001000
J_Q05bJPX			6	10 years or more	0000100
J_Q05bJPX			96	Valid skip	0000010
J_Q05bKO	9	KO_Background - Language spoken at home	-1	Missing	00000001
J_Q05bKO			1	Korean	00000000
J_Q05bKO			2	Chinese	10000000
J_Q05bKO			3	English	01000000
J_Q05bKO			4	Vietnamese	00100000
J_Q05bKO			5	Filipino	00010000
J_Q05bKO			6	Japanese	00001000
J_Q05bKO			7	Other language	00000100
J_Q05bKO			96	Valid skip	00000010
J_Q05bNL	9	Background - Language spoken at home	-1	Missing	00000001
J_Q05bNL			1	dutch	00000000
J_Q05bNL			2	arabic	10000000
J_Q05bNL			3	turkish	01000000
J_Q05bNL			4	chinese	00100000
J_Q05bNL			5	french	00010000
J_Q05bNL			6	english	00001000
J_Q05bNL			7	other language	00000100
J_Q05bNL			96	Valid skip	00000010
J_Q05bNO	21	Background - Language spoken at home	-1	Missing	00000000000000000001
J_Q05bNO			1	Norwegian	00000000000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05bNO			2	Danish	10000000000000000000
J_Q05bNO			3	English	01000000000000000000
J_Q05bNO			4	French	00100000000000000000
J_Q05bNO			5	Hindi	00010000000000000000
J_Q05bNO			6	Kurd	00001000000000000000
J_Q05bNO			7	Persian	00000100000000000000
J_Q05bNO			8	Punjabi	00000010000000000000
J_Q05bNO			9	Serbian	00000001000000000000
J_Q05bNO			10	Serbo-Croat	00000000100000000000
J_Q05bNO			11	Singhalese	00000000010000000000
J_Q05bNO			12	Somali	00000000001000000000
J_Q05bNO			13	Spanish	00000000000100000000
J_Q05bNO			14	Swedish	00000000000010000000
J_Q05bNO			15	Turkish	00000000000001000000
J_Q05bNO			16	German	00000000000000100000
J_Q05bNO			17	Urdu	00000000000000010000
J_Q05bNO			18	Vietnamese	00000000000000001000
J_Q05bNO			19	Other language	00000000000000000100
J_Q05bNO			96	Valid skip	00000000000000000010
J_Q05bPL	14	Background - Language spoken at home	-1	Missing	0000000000001
J_Q05bPL			1	Byelorussian	0000000000000
J_Q05bPL			2	Czech	1000000000000
J_Q05bPL			3	Dutch	0100000000000
J_Q05bPL			4	English	0010000000000
J_Q05bPL			5	French	0001000000000
J_Q05bPL			6	German	0000100000000
J_Q05bPL			7	Lithuanian	0000010000000
J_Q05bPL			8	Polish	0000001000000
J_Q05bPL			9	Russian	0000000100000
J_Q05bPL			10	Slovak	0000000010000
J_Q05bPL			11	Ukrainian	0000000001000
J_Q05bPL			12	Other language	0000000000100
J_Q05bPL			96	Valid skip	0000000000010
J_Q05bRU	9	Background - Language spoken at home	-1	Missing	00000001
J_Q05bRU			1	Language1	00000000
J_Q05bRU			2	Language2	10000000
J_Q05bRU			3	Language3	01000000
J_Q05bRU			4	Language4	00100000
J_Q05bRU			5	Language5	00010000
J_Q05bRU			6	Language6	00001000
J_Q05bRU			7	Other language	00000100
J_Q05bRU			96	Valid skip	00000010
J_Q05bSE	14	Background - Language spoken at home	-1	Missing	0000000000001
J_Q05bSE			1	Svenska	0000000000000
J_Q05bSE			2	Finska	1000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05bSE			3	Spanska	0100000000000
J_Q05bSE			4	Arabiska	0010000000000
J_Q05bSE			5	Persiska	0001000000000
J_Q05bSE			6	Polska	0000100000000
J_Q05bSE			7	Serbokroatiska	0000010000000
J_Q05bSE			8	Engelska	0000001000000
J_Q05bSE			9	Turkiska	0000000100000
J_Q05bSE			10	Bosniska	0000000010000
J_Q05bSE			11	Kurdiska	0000000001000
J_Q05bSE			12	Kurdiska k ange	0000000000100
J_Q05bSE			96	Valid skip	0000000000010
J_Q05bSK	9	Background - Language spoken at home	-1	Missing	00000001
J_Q05bSK			1	Slovak	00000000
J_Q05bSK			2	Czech	10000000
J_Q05bSK			3	Hungarian	01000000
J_Q05bSK			4	German	00100000
J_Q05bSK			5	Roma	00010000
J_Q05bSK			6	Polish	00001000
J_Q05bSK			7	Other language	00000100
J_Q05bSK			96	Valid skip	00000010
J_Q05bUK	12	Background - Language spoken at home	-1	Missing	00000000001
J_Q05bUK			1	English	00000000000
J_Q05bUK			2	Welsh	10000000000
J_Q05bUK			3	Irish	01000000000
J_Q05bUK			4	Scottish Gaelic	00100000000
J_Q05bUK			5	Ulster Scots/Ullans	00010000000
J_Q05bUK			6	Hindi	00001000000
J_Q05bUK			7	Urdu	00000100000
J_Q05bUK			8	Punjabi	00000010000
J_Q05bUK			9	Polish	00000001000
J_Q05bUK			10	Other	00000000100
J_Q05bUK			96	Valid skip	00000000010
J_Q05bUS	9	Background - Language spoken at home	-1	Missing	00000001
J_Q05bUS			1	English	00000000
J_Q05bUS			2	Spanis	10000000
J_Q05bUS			3	French	01000000
J_Q05bUS			4	Italian	00100000
J_Q05bUS			5	Chinese	00010000
J_Q05bUS			6	German	00001000
J_Q05bUS			7	Other language	00000100
J_Q05bUS			96	Valid skip	00000010
J_Q05cUSX1	7	Background - Language spoken most	-1	Missing	000001
J_Q05cUSX1			1	English only	000000
J_Q05cUSX1			2	English and Spanish	100000
J_Q05cUSX1			3	English and Other	010000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q05cUSX1			4	Spanish only	001000
J_Q05cUSX1			5	Other only	000100
J_Q05cUSX1			6	Valid skip	000010
J_Q05cUSX2	4	Background - English outside home	-1	Missing	001
J_Q05cUSX2			1	Yes	000
J_Q05cUSX2			2	No	100
J_Q05cUSX2			6	Valid skip	010
J_Q05cUSX3a	6	Background - Ability to understand spoken English	-1	Missing	00001
J_Q05cUSX3a			1	Very well	00000
J_Q05cUSX3a			2	Well	10000
J_Q05cUSX3a			3	Not well	01000
J_Q05cUSX3a			4	Not at all	00100
J_Q05cUSX3a			6	Valid skip	00010
J_Q05cUSX3b	6	Background - Ability to speak English	-1	Missing	00001
J_Q05cUSX3b			1	Very well	00000
J_Q05cUSX3b			2	Well	10000
J_Q05cUSX3b			3	Not well	01000
J_Q05cUSX3b			4	Not at all	00100
J_Q05cUSX3b			6	Valid skip	00010
J_Q05cUSX3d	6	Background - Ability to read English	-1	Missing	00001
J_Q05cUSX3d			1	Very well	00000
J_Q05cUSX3d			2	Well	10000
J_Q05cUSX3d			3	Not well	01000
J_Q05cUSX3d			4	Not at all	00100
J_Q05cUSX3d			6	Valid skip	00010
J_Q05cUSX3e	6	Background - Ability to write English	-1	Missing	00001
J_Q05cUSX3e			1	Very well	00000
J_Q05cUSX3e			2	Well	10000
J_Q05cUSX3e			3	Not well	01000
J_Q05cUSX3e			4	Not at all	00100
J_Q05cUSX3e			6	Valid skip	00010
J_Q05cUSX4	4	Background - ESL class/tutor in past year	-1	Missing	001
J_Q05cUSX4			1	Yes	000
J_Q05cUSX4			2	No	100
J_Q05cUSX4			6	Valid skip	010
J_Q05cUSX5	5	Background - Reason for ESL class/tutor	-1	Missing	0001
J_Q05cUSX5			1	WORK-RELATED	0000
J_Q05cUSX5			2	PERSONAL INTEREST	1000
J_Q05cUSX5			3	BOTH EQUALLY	0100
J_Q05cUSX5			6	Valid skip	0010
J_Q05cUSX6	4	Background - Class/tutor learn English as adult	-1	Missing	001
J_Q05cUSX6			1	Yes	000
J_Q05cUSX6			2	No	100
J_Q05cUSX6			6	Valid skip	010
J_Q06aAU	4	Background - Mother/female guardian - Whether born	-1	Missing	001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q06aAU			1	Yes	000
J_Q06aAU			2	No	100
J_Q06aAU			6	Valid skip	010
J_Q06aDEX	11	Background - Mother/female guardian - country of b	-1	Missing	0000000001
J_Q06aDEX			1	Turkey	0000000000
J_Q06aDEX			2	Italy	1000000000
J_Q06aDEX			3	Poland	0100000000
J_Q06aDEX			4	Greece	0010000000
J_Q06aDEX			5	Serbia	0001000000
J_Q06aDEX			6	Croatia	0000100000
J_Q06aDEX			7	Russian Federation	0000010000
J_Q06aDEX			8	Bosnia and Herzegovi	0000001000
J_Q06aDEX			9	Another country	0000000100
J_Q06aDEX			96	Valid skip	0000000010
J_Q06bAT	9	Background - Mother/female guardian - Highest leve	-1	Missing	00000001
J_Q06bAT			1	Compulsory school	00000000
J_Q06bAT			2	Apprenticeship	10000000
J_Q06bAT			3	Vocational School	01000000
J_Q06bAT			4	Master Craftsman's c	00100000
J_Q06bAT			5	Secondary school wit	00010000
J_Q06bAT			6	Academic Study	00001000
J_Q06bAT			7	Other education afte	00000100
J_Q06bAT			96	Valid skip	00000010
J_Q06bAU	17	Background - Mother/female guardian - Highest leve	-1	Missing	0000000000000001
J_Q06bAU			1	Year 8 or below	0000000000000000
J_Q06bAU			2	Year 9 or equivalent	1000000000000000
J_Q06bAU			3	Year 10 or equivalen	0100000000000000
J_Q06bAU			4	Year 11 or equivalen	0010000000000000
J_Q06bAU			5	Year 12 or equivalen	0001000000000000
J_Q06bAU			6	Certificate I	0000100000000000
J_Q06bAU			7	Certificate II	0000010000000000
J_Q06bAU			8	Certificate III	0000001000000000
J_Q06bAU			9	Certificate IV	0000000100000000
J_Q06bAU			10	Diploma	0000000010000000
J_Q06bAU			11	Advanced Diploma and	0000000001000000
J_Q06bAU			12	Bachelor degree (inc	0000000000100000
J_Q06bAU			13	Graduate Diploma or	0000000000010000
J_Q06bAU			14	Masters	0000000000001000
J_Q06bAU			15	Doctorate	0000000000000100
J_Q06bAU			96	Valid skip	0000000000000010
J_Q06bCA	9	Background - Mother/female guardian - Highest leve	-1	Missing	00000001
J_Q06bCA			1	No formal education	00000000
J_Q06bCA			2	Less than high schoo	10000000
J_Q06bCA			3	High school diploma	01000000
J_Q06bCA			4	Apprenticeship certi	00100000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q06bCA			5	Trade/vocational cer	00010000
J_Q06bCA			6	Non-university certi	00001000
J_Q06bCA			7	University certifica	00000100
J_Q06bCA			96	Valid skip	00000010
J_Q06bCZ	7	Background - Mother/female guardian - Highest leve	-1	Missing	000001
J_Q06bCZ			1	ISCED123cshort basic	000000
J_Q06bCZ			2	ISCED3C vocational l	100000
J_Q06bCZ			3	ISCED 3A upper secon	010000
J_Q06bCZ			4	ISCED 4, 5B post sec	001000
J_Q06bCZ			5	ISCED5A, 6 universit	000100
J_Q06bCZ			6	Valid skip	000010
J_Q06bDE1	10	Background - Mother/female guardian - Highest leve	-1	Missing	000000001
J_Q06bDE1			1	Left school without	000000000
J_Q06bDE1			2	Hauptschulabschluss	100000000
J_Q06bDE1			3	Realschulabschluss (	010000000
J_Q06bDE1			4	Left the Polytechnis	001000000
J_Q06bDE1			5	Left the Polytechnis	000100000
J_Q06bDE1			6	Fachhochschulereife,	000010000
J_Q06bDE1			7	Abitur/EOS (General	000001000
J_Q06bDE1			8	Another school leavi	000000100
J_Q06bDE1			96	Valid skip	000000010
J_Q06bDE1_REC	10	Background - Mother/female guardian - Highest leve	-1	Missing	000000001
J_Q06bDE1_REC			1	Left school without	000000000
J_Q06bDE1_REC			2	Hauptschulabschluss	100000000
J_Q06bDE1_REC			3	Realschulabschluss (	010000000
J_Q06bDE1_REC			4	Left the Polytechnis	001000000
J_Q06bDE1_REC			5	Left the Polytechnis	000100000
J_Q06bDE1_REC			6	Fachhochschulereife,	000010000
J_Q06bDE1_REC			7	Abitur/EOS (General	000001000
J_Q06bDE1_REC			8	Another school leavi	000000100
J_Q06bDE1_REC			96	Valid skip	000000010
J_Q06bDE2	13	Background - Mother/female guardian - Highest leve	-1	Missing	000000000001
J_Q06bDE2			1	No professional qual	000000000000
J_Q06bDE2			2	Apprenticeship (Lehr	100000000000
J_Q06bDE2			3	Basic vocational tra	010000000000
J_Q06bDE2			4	Training at Fachschu	001000000000
J_Q06bDE2			5	Berufsakademie, Fach	000100000000
J_Q06bDE2			6	Bachelor at Fachhoch	000010000000
J_Q06bDE2			7	Master/Diplom at Fac	000001000000
J_Q06bDE2			8	Bachelor at universi	000000100000
J_Q06bDE2			9	Master/Diplom at uni	000000010000
J_Q06bDE2			10	Doctorate	000000001000
J_Q06bDE2			11	Another professional	000000000100
J_Q06bDE2			96	Valid skip	000000000010
J_Q06bDE2_REC	13	Background - Mother/female guardian - Highest leve	-1	Missing	000000000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q06bDE2_REC			1	No professional qual	000000000000
J_Q06bDE2_REC			2	Apprenticeship (Lehr	100000000000
J_Q06bDE2_REC			3	Basic vocational tra	010000000000
J_Q06bDE2_REC			4	Training at Fachschu	001000000000
J_Q06bDE2_REC			5	Berufsakademie, Fach	000100000000
J_Q06bDE2_REC			6	Bachelor at Fachhoch	000010000000
J_Q06bDE2_REC			7	Master/Diplom at Fac	000001000000
J_Q06bDE2_REC			8	Bachelor at universi	000000100000
J_Q06bDE2_REC			9	Master/Diplom at uni	000000010000
J_Q06bDE2_REC			10	Doctorate	000000001000
J_Q06bDE2_REC			11	Another professional	000000000100
J_Q06bDE2_REC			96	Valid skip	000000000010
J_Q06bFR	16	Background - Mother/female guardian - Highest leve	-1	Missing	000000000000001
J_Q06bFR			1	No formal qualificat	000000000000000
J_Q06bFR			2	ISCED 1	100000000000000
J_Q06bFR			3	ISCED 2	010000000000000
J_Q06bFR			4	ISCED 3C shorter tha	001000000000000
J_Q06bFR			5	ISCED 3C 2 years or	000100000000000
J_Q06bFR			6	ISCED 3A-B	000010000000000
J_Q06bFR			7	ISCED 3 (without dis	000001000000000
J_Q06bFR			8	ISCED 4C	000000100000000
J_Q06bFR			9	ISCED 4A-B	000000010000000
J_Q06bFR			10	ISCED 4 (without dis	000000001000000
J_Q06bFR			11	ISCED 5B	000000000100000
J_Q06bFR			12	ISCED 5A, bachelor d	000000000010000
J_Q06bFR			13	ISCED 5A, master deg	000000000001000
J_Q06bFR			14	ISCED 6	000000000000100
J_Q06bFR			96	Valid skip	000000000000010
J_Q06bPL	6	Background - Mother/female guardian - Highest leve	-1	Missing	00001
J_Q06bPL			1	ISCED123cshort	00000
J_Q06bPL			2	ISCED3clong	10000
J_Q06bPL			3	ISCED3ba4	01000
J_Q06bPL			4	ISCED56	00100
J_Q06bPL			6	Valid skip	00010
J_Q06bUK	11	Background - Mother/female guardian - Highest leve	-1	Missing	0000000001
J_Q06bUK			1	No qualifications	0000000000
J_Q06bUK			2	Key Skills, Basic sk	1000000000
J_Q06bUK			3	O levels, GCSE or eq	0100000000
J_Q06bUK			4	NVQ Level2, City & G	0010000000
J_Q06bUK			5	A Levels or equivale	0001000000
J_Q06bUK			6	Trade apprenticeship	0000100000
J_Q06bUK			7	NVQ Level 3, City &	0000010000
J_Q06bUK			8	Degree or higher deg	0000001000
J_Q06bUK			9	NVQ Level 4 or 5, HN	0000000100
J_Q06bUK			96	Valid skip	0000000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q06bUS	5	Background - Mother/female guardian - Highest leve	-1	Missing	0001
J_Q06bUS			1	Less than high schoo	0000
J_Q06bUS			2	High school diploma/	1000
J_Q06bUS			3	College degree or hi	0100
J_Q06bUS			6	Valid skip	0010
J_Q06cBE	5	Mother/female guardian - paid job	-1	Missing	0001
J_Q06cBE			1	Yes	0000
J_Q06cBE			2	No	1000
J_Q06cBE			3	N/A	0100
J_Q06cBE			6	Valid skip	0010
J_Q06cCZ	5	Background - Mother/female guardien - paid job	-1	Missing	0001
J_Q06cCZ			1	Yes	0000
J_Q06cCZ			2	No	1000
J_Q06cCZ			3	Not applicable, no f	0100
J_Q06cCZ			96	Valid skip	0010
J_Q06cDEX	5	Background - Mother - Hold a paying job	-1	Missing	0001
J_Q06cDEX			1	Yes	0000
J_Q06cDEX			2	No	1000
J_Q06cDEX			3	Not applicable, moth	0100
J_Q06cDEX			6	Valid skip	0010
J_Q06cES	5	Empleo remunerado madre o tutora	-1	Missing	0001
J_Q06cES			1	Not stated or inferr	0000
J_Q06cES			2	No	1000
J_Q06cES			3	No es pertinente, ma	0100
J_Q06cES			6	Valid skip	0010
J_Q06cIE	5	Background - Mother\female guardian - Work situati	-1	Missing	0001
J_Q06cIE			1	Yes	0000
J_Q06cIE			2	No	1000
J_Q06cIE			3	Not applicable, pare	0100
J_Q06cIE			96	Valid skip	0010
J_Q06cIT	5	Background - Mother/female guardian - Hold a payin	-1	Missing	0001
J_Q06cIT			1	Yes	0000
J_Q06cIT			2	No	1000
J_Q06cIT			3	Not applicable, no m	0100
J_Q06cIT			6	Valid skip	0010
J_Q06cPL	5	Background - Mother/female guardian - Hold a payin	-1	Missing	0001
J_Q06cPL			1	Yes	0000
J_Q06cPL			2	No	1000
J_Q06cPL			3	Not applicable, no m	0100
J_Q06cPL			6	Valid skip	0010
J_Q06cUK	5	Mother/Female guardian - hold paying job	-1	Missing	0001
J_Q06cUK			1	Yes	0000
J_Q06cUK			2	No	1000
J_Q06cUK			3	Not applicable, moth	0100
J_Q06cUK			6	Valid skip	0010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q06dFR1	6	Background - Mother/female guardian - Job status	-1	Missing	00001
J_Q06dFR1			1	Running his/her own	00000
J_Q06dFR1			2	Helping one of his/h	10000
J_Q06dFR1			3	As a civil servant w	01000
J_Q06dFR1			4	As an employee	00100
J_Q06dFR1			6	Valid skip	00010
J_Q06eFR	12	Background - Mother/female guardian - Job main tas	-1	Missing	00000000001
J_Q06eFR			1	Production, construc	00000000000
J_Q06eFR			2	Repairing, maintaini	10000000000
J_Q06eFR			3	Cleaning, caretaking	01000000000
J_Q06eFR			4	Handing, logistics	00100000000
J_Q06eFR			5	Secretary, reception	00010000000
J_Q06eFR			6	Accounting, administ	00001000000
J_Q06eFR			7	Sales and marketing	00000100000
J_Q06eFR			8	Research and develop	00000010000
J_Q06eFR			9	Education, healthcar	00000001000
J_Q06eFR			10	Other. Specify.	00000000100
J_Q06eFR			96	Valid skip	00000000010
J_Q07aAU	4	Background - Father/male guardian - Whether born i	-1	Missing	001
J_Q07aAU			1	Yes	000
J_Q07aAU			2	No	100
J_Q07aAU			6	Valid skip	010
J_Q07aDEX	11	Background - Father/male guardian - country of bir	-1	Missing	0000000001
J_Q07aDEX			1	Turkey	0000000000
J_Q07aDEX			2	Italy	1000000000
J_Q07aDEX			3	Poland	0100000000
J_Q07aDEX			4	Greece	0010000000
J_Q07aDEX			5	Serbia	0001000000
J_Q07aDEX			6	Croatia	0000100000
J_Q07aDEX			7	Russian Federation	0000010000
J_Q07aDEX			8	Bosnia and Herzegovi	0000001000
J_Q07aDEX			9	Another country	0000000100
J_Q07aDEX			96	Valid skip	0000000010
J_Q07bAT	9	Background - Father/male guardian - Highest level	-1	Missing	00000001
J_Q07bAT			1	Compulsory school	00000000
J_Q07bAT			2	Apprenticeship	10000000
J_Q07bAT			3	Vocational School	01000000
J_Q07bAT			4	Master Craftsman's c	00100000
J_Q07bAT			5	Secondary school wit	00010000
J_Q07bAT			6	Academic Study	00001000
J_Q07bAT			7	Other education afte	00000100
J_Q07bAT			96	Valid skip	00000010
J_Q07bAU	17	Background - Father/male guardian - Highest level	-1	Missing	0000000000000001
J_Q07bAU			1	Year 8 or below	0000000000000000
J_Q07bAU			2	Year 9 or equivalent	1000000000000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q07bAU			3	Year 10 or equivalen	0100000000000000
J_Q07bAU			4	Year 11 or equivalen	0010000000000000
J_Q07bAU			5	Year 12 or equivalen	0001000000000000
J_Q07bAU			6	Certificate I	0000100000000000
J_Q07bAU			7	Certificate II	0000010000000000
J_Q07bAU			8	Certificate III	0000001000000000
J_Q07bAU			9	Certificate IV	0000000100000000
J_Q07bAU			10	Diploma	0000000010000000
J_Q07bAU			11	Advanced Diploma and	0000000001000000
J_Q07bAU			12	Bachelor degree (inc	0000000000100000
J_Q07bAU			13	Graduate Diploma or	0000000000010000
J_Q07bAU			14	Masters	0000000000001000
J_Q07bAU			15	Doctorate	0000000000000100
J_Q07bAU			96	Valid skip	0000000000000010
J_Q07bCA	9	Background - Father/male guardian - Highest level	-1	Missing	00000001
J_Q07bCA			1	No formal education	00000000
J_Q07bCA			2	Less than high schoo	10000000
J_Q07bCA			3	High school diploma	01000000
J_Q07bCA			4	Apprenticeship certi	00100000
J_Q07bCA			5	Trade/vocational cer	00010000
J_Q07bCA			6	Non-university certi	00001000
J_Q07bCA			7	University certifica	00000100
J_Q07bCA			96	Valid skip	00000010
J_Q07bCZ	7	Background - Father/male guardian - Highest level	-1	Missing	000001
J_Q07bCZ			1	ISCED123cshort basic	000000
J_Q07bCZ			2	ISCED3C vocational l	100000
J_Q07bCZ			3	ISCED 3A upper secon	010000
J_Q07bCZ			4	ISCED 4, 5B post sec	001000
J_Q07bCZ			5	ISCED5A, 6 universit	000100
J_Q07bCZ			6	Valid skip	000010
J_Q07bDE1	10	Background - Father/male guardian - Highest level	-1	Missing	000000001
J_Q07bDE1			1	Left school without	000000000
J_Q07bDE1			2	Hauptschulabschluss	100000000
J_Q07bDE1			3	Realschulabschluss (	010000000
J_Q07bDE1			4	Left the Polytechnis	001000000
J_Q07bDE1			5	Left the Polytechnis	000100000
J_Q07bDE1			6	Fachhochschulereife,	000010000
J_Q07bDE1			7	Abitur/EOS (General	000001000
J_Q07bDE1			8	Another school leavi	000000100
J_Q07bDE1			96	Valid skip	000000010
J_Q07bDE1_REC	10	Background - Father/male guardian - Highest level	-1	Missing	000000001
J_Q07bDE1_REC			1	Left school without	000000000
J_Q07bDE1_REC			2	Hauptschulabschluss	100000000
J_Q07bDE1_REC			3	Realschulabschluss (	010000000
J_Q07bDE1_REC			4	Left the Polytechnis	001000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q07bDE1_REC			5	Left the Polytechnis	000100000
J_Q07bDE1_REC			6	Fachhochschulereife,	000010000
J_Q07bDE1_REC			7	Abitur/EOS (General	000001000
J_Q07bDE1_REC			8	Another school leavi	000000100
J_Q07bDE1_REC			96	Valid skip	000000010
J_Q07bDE2	13	Background - Father/male guardian - Highest level	-1	Missing	000000000001
J_Q07bDE2			1	No professional qual	000000000000
J_Q07bDE2			2	Apprenticeship (Lehr	100000000000
J_Q07bDE2			3	Basic vocational tra	010000000000
J_Q07bDE2			4	Training at Fachschu	001000000000
J_Q07bDE2			5	Berufsakademie, Fach	000100000000
J_Q07bDE2			6	Bachelor at Fachhoch	000010000000
J_Q07bDE2			7	Master/Diplom at Fac	000001000000
J_Q07bDE2			8	Bachelor at universi	000000100000
J_Q07bDE2			9	Master/Diplom at uni	000000010000
J_Q07bDE2			10	Doctorate	000000001000
J_Q07bDE2			11	Another professional	000000000100
J_Q07bDE2			96	Valid skip	000000000010
J_Q07bDE2_REC	13	Background - Father/male guardian - Highest level	-1	Missing	000000000001
J_Q07bDE2_REC			1	No professional qual	000000000000
J_Q07bDE2_REC			2	Apprenticeship (Lehr	100000000000
J_Q07bDE2_REC			3	Basic vocational tra	010000000000
J_Q07bDE2_REC			4	Training at Fachschu	001000000000
J_Q07bDE2_REC			5	Berufsakademie, Fach	000100000000
J_Q07bDE2_REC			6	Bachelor at Fachhoch	000010000000
J_Q07bDE2_REC			7	Master/Diplom at Fac	000001000000
J_Q07bDE2_REC			8	Bachelor at universi	000000100000
J_Q07bDE2_REC			9	Master/Diplom at uni	000000010000
J_Q07bDE2_REC			10	Doctorate	000000001000
J_Q07bDE2_REC			11	Another professional	000000000100
J_Q07bDE2_REC			96	Valid skip	000000000010
J_Q07bFR	16	Background - Father/male guardian - Highest level	-1	Missing	000000000000001
J_Q07bFR			1	No formal qualificat	000000000000000
J_Q07bFR			2	ISCED 1	100000000000000
J_Q07bFR			3	ISCED 2	010000000000000
J_Q07bFR			4	ISCED 3C shorter tha	001000000000000
J_Q07bFR			5	ISCED 3C 2 years or	000100000000000
J_Q07bFR			6	ISCED 3A-B	000010000000000
J_Q07bFR			7	ISCED 3 (without dis	000001000000000
J_Q07bFR			8	ISCED 4C	000000100000000
J_Q07bFR			9	ISCED 4A-B	000000010000000
J_Q07bFR			10	ISCED 4 (without dis	000000001000000
J_Q07bFR			11	ISCED 5B	000000000100000
J_Q07bFR			12	ISCED 5A, bachelor d	000000000010000
J_Q07bFR			13	ISCED 5A, master deg	000000000001000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q07bFR			14	ISCED 6	000000000000100
J_Q07bFR			96	Valid skip	000000000000010
J_Q07bPL	6	Background - Father/male guardian - Highest level	-1	Missing	00001
J_Q07bPL			1	ISCED123cshort	00000
J_Q07bPL			2	ISCED3clong	10000
J_Q07bPL			3	ISCED3ba4	01000
J_Q07bPL			4	ISCED56	00100
J_Q07bPL			6	Valid skip	00010
J_Q07bUK	11	Background - Father/male guardian - Highest level	-1	Missing	0000000001
J_Q07bUK			1	No qualifications	0000000000
J_Q07bUK			2	Key Skills, Basic sk	1000000000
J_Q07bUK			3	O levels, GCSE or eq	0100000000
J_Q07bUK			4	NVQ Level2, City & G	0010000000
J_Q07bUK			5	A Levels or equivale	0001000000
J_Q07bUK			6	Trade apprenticeship	0000100000
J_Q07bUK			7	NVQ Level 3, City &	0000010000
J_Q07bUK			8	Degree or higher deg	0000001000
J_Q07bUK			9	NVQ Level 4 or 5, HN	0000000100
J_Q07bUK			96	Valid skip	0000000010
J_Q07bUS	5	Background - Father/male guardian - Highest level	-1	Missing	0001
J_Q07bUS			1	Less than high schoo	0000
J_Q07bUS			2	High school diploma/	1000
J_Q07bUS			3	College degree or hi	0100
J_Q07bUS			6	Valid skip	0010
J_Q07cBE	5	Paying job - Father/male guardian	-1	Missing	0001
J_Q07cBE			1	Yes	0000
J_Q07cBE			2	No	1000
J_Q07cBE			3	N/A	0100
J_Q07cBE			6	Valid skip	0010
J_Q07cCZ	5	Background - Father/male guardien - paid job	-1	Missing	0001
J_Q07cCZ			1	Yes	0000
J_Q07cCZ			2	No	1000
J_Q07cCZ			3	Not applicable, no f	0100
J_Q07cCZ			96	Valid skip	0010
J_Q07cDEX	5	Background - Father - Hold a paying job	-1	Missing	0001
J_Q07cDEX			1	Yes	0000
J_Q07cDEX			2	No	1000
J_Q07cDEX			3	Not applicable, moth	0100
J_Q07cDEX			6	Valid skip	0010
J_Q07cES	5	Empleo remunerado padre o tutor	-1	Missing	0001
J_Q07cES			1	Not stated or inferr	0000
J_Q07cES			2	No	1000
J_Q07cES			3	No es pertinente, pa	0100
J_Q07cES			6	Valid skip	0010
J_Q07cIE	5	Background - Father\male guardian - Work situation	-1	Missing	0001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q07cIE			1	Yes	0000
J_Q07cIE			2	No	1000
J_Q07cIE			3	Not applicable, pare	0100
J_Q07cIE			96	Valid skip	0010
J_Q07cIT	5	Background - Father/male guardian - Hold a paying	-1	Missing	0001
J_Q07cIT			1	Yes	0000
J_Q07cIT			2	No	1000
J_Q07cIT			3	Not applicable, no f	0100
J_Q07cIT			6	Valid skip	0010
J_Q07cPL	5	Background - Father/male guardian - Hold a paying	-1	Missing	0001
J_Q07cPL			1	Yes	0000
J_Q07cPL			2	No	1000
J_Q07cPL			3	Not applicable, no f	0100
J_Q07cPL			6	Valid skip	0010
J_Q07cUK	5	Father/male guardian - hold a paying job	-1	Missing	0001
J_Q07cUK			1	Yes	0000
J_Q07cUK			2	No	1000
J_Q07cUK			3	Not applicable, fath	0100
J_Q07cUK			6	Valid skip	0010
J_Q07dFR1	6	Background - Father/male guardian - Job status	-1	Missing	00001
J_Q07dFR1			1	Running his/her own	00000
J_Q07dFR1			2	Helping one of his/h	10000
J_Q07dFR1			3	As a civil servant w	01000
J_Q07dFR1			4	As an employee	00100
J_Q07dFR1			6	Valid skip	00010
J_Q07eFR	12	Background - Father/male guardian - Job main task	-1	Missing	00000000001
J_Q07eFR			1	Production, construc	00000000000
J_Q07eFR			2	Repairing, maintaini	10000000000
J_Q07eFR			3	Cleaning, caretaking	01000000000
J_Q07eFR			4	Handing, logistics	00100000000
J_Q07eFR			5	Secretary, reception	00010000000
J_Q07eFR			6	Accounting, administ	00001000000
J_Q07eFR			7	Sales and marketing	00000100000
J_Q07eFR			8	Research and develop	00000010000
J_Q07eFR			9	Education, healthcar	00000001000
J_Q07eFR			10	Other. Specify.	00000000100
J_Q07eFR			96	Valid skip	00000000010
J_Q08AU	8	Background - Number of books at home (AUS)	-1	Missing	0000001
J_Q08AU			1	10 books or less	0000000
J_Q08AU			2	11 to 25 books	1000000
J_Q08AU			3	26 to 100 books	0100000
J_Q08AU			4	101 to 200 books	0010000
J_Q08AU			5	201 to 500 books	0001000
J_Q08AU			6	More than 500 books	0000100
J_Q08AU			96	Valid skip	0000010

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q09aIEX	4	Income sources - Unemployment benefit	-1	Missing	001
J_Q09aIEX			1	Yes	000
J_Q09aIEX			2	No	100
J_Q09aIEX			6	Valid skip	010
J_Q09bIEX	4	Income sources - Disability benefit	-1	Missing	001
J_Q09bIEX			1	Yes	000
J_Q09bIEX			2	No	100
J_Q09bIEX			6	Valid skip	010
J_Q09cIEX	4	Income sources - Illness benefit	-1	Missing	001
J_Q09cIEX			1	Yes	000
J_Q09cIEX			2	No	100
J_Q09cIEX			6	Valid skip	010
J_Q09dIEX	4	Income sources - Early retirement benefit	-1	Missing	001
J_Q09dIEX			1	Yes	000
J_Q09dIEX			2	No	100
J_Q09dIEX			6	Valid skip	010
J_q09edkx2	4	Do you expect to stop working entirely when you re	-1	Missing	001
J_q09edkx2			1	I expect to stop wor	000
J_q09edkx2			2	I expect to retire g	100
J_q09edkx2			96	Valid skip	010
J_Q09edkx3	5	Do you expect to retire because you have to.....?	-1	Missing	0001
J_Q09edkx3			1	I expect to retire w	0000
J_Q09edkx3			2	I expect that retire	1000
J_Q09edkx3			3	I expect	0100
J_Q09edkx3			96	Valid skip	0010
J_Q09edkx4	7	What do you expect will be the primary source of i	-1	Missing	000001
J_Q09edkx4			1	Early retirement wag	000000
J_Q09edkx4			2	Own pension savings	100000
J_Q09edkx4			3	Old age pension	010000
J_Q09edkx4			4	Disability pension	001000
J_Q09edkx4			5	Other	000100
J_Q09edkx4			96	Valid skip	000010
J_Q09eIEX	4	Income sources - Retirement benefit	-1	Missing	001
J_Q09eIEX			1	Yes	000
J_Q09eIEX			2	No	100
J_Q09eIEX			6	Valid skip	010
J_Q09eIEX1	4	Income sources - Maternity benefit	-1	Missing	001
J_Q09eIEX1			1	Yes	000
J_Q09eIEX1			2	No	100
J_Q09eIEX1			6	Valid skip	010
J_Q09eIEX2	4	Income sources - Family Income Supplement	-1	Missing	001
J_Q09eIEX2			1	Yes	000
J_Q09eIEX2			2	No	100
J_Q09eIEX2			6	Valid skip	010
J_Q10UKX1	13	Background - Religion - Scotland	-1	Missing	000000000001

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
J_Q10UKX1			1	None	000000000000
J_Q10UKX1			2	Church of Scotland	100000000000
J_Q10UKX1			3	Roman Catholic	010000000000
J_Q10UKX1			4	Other Christian	001000000000
J_Q10UKX1			5	Muslim	000100000000
J_Q10UKX1			6	Buddhist	000010000000
J_Q10UKX1			7	Sikh	000001000000
J_Q10UKX1			8	Jewish	000000100000
J_Q10UKX1			9	Hindu	000000010000
J_Q10UKX1			10	Pagan	000000001000
J_Q10UKX1			11	Another Religion	000000000100
J_Q10UKX1			96	Valid skip	000000000010
J_Q10UKX2	19	Background - Religion - NI	-1	Missing	0000000000000000001
J_Q10UKX2			1	Catholic	0000000000000000000
J_Q10UKX2			2	Presbyterian	1000000000000000000
J_Q10UKX2			3	Church of Ireland	0100000000000000000
J_Q10UKX2			4	Methodist	0010000000000000000
J_Q10UKX2			5	Baptist	0001000000000000000
J_Q10UKX2			6	Free Presbyterian	0000100000000000000
J_Q10UKX2			7	Brethren	0000010000000000000
J_Q10UKX2			8	Protestant - not spe	0000001000000000000
J_Q10UKX2			9	Other Christian	0000000100000000000
J_Q10UKX2			10	Buddhist	0000000010000000000
J_Q10UKX2			11	Hindu	0000000001000000000
J_Q10UKX2			12	Jewish	0000000000100000000
J_Q10UKX2			13	Muslim	0000000000010000000
J_Q10UKX2			14	Sikh	0000000000001000000
J_Q10UKX2			15	Other Religion	0000000000000100000
J_Q10UKX2			16	Unwilling to answer	0000000000000010000
J_Q10UKX2			17	No religion	0000000000000001000
J_Q10UKX2			96	Valid skip	0000000000000000010
J_Q10UKX3	11	Background - Religion - England	-1	Missing	0000000001
J_Q10UKX3			1	Christian (inc CoE,	0000000000
J_Q10UKX3			2	Buddhist	1000000000
J_Q10UKX3			3	Hindu	0100000000
J_Q10UKX3			4	Jewish	0010000000
J_Q10UKX3			5	Muslim	0001000000
J_Q10UKX3			6	Sikh	0000100000
J_Q10UKX3			7	Other religion	0000010000
J_Q10UKX3			8	Unwilling to answer	0000001000
J_Q10UKX3			9	No religion	0000000100
J_Q10UKX3			96	Valid skip	0000000010
K_Q01AU	4	Income - Wages or salaries	-1	Missing	001
K_Q01AU			1	Yes	000
K_Q01AU			2	No	100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
K_Q01AU			6	Valid skip	010
K_Q01bAU	8	Income - Wages or salaries - Period	-1	Missing	0000001
K_Q01bAU			1	Week	0000000
K_Q01bAU			2	Fortnight	1000000
K_Q01bAU			3	Four weeks	0100000
K_Q01bAU			4	Calendar month	0010000
K_Q01bAU			5	Year	0001000
K_Q01bAU			6	Other (please specif	0000100
K_Q01bAU			96	Valid skip	0000010
K_Q02bAU	14	Income - Current pensions	-1	Missing	0000000000001
K_Q02bAU			1	Australian Age Pensi	0000000000000
K_Q02bAU			2	Service Pension from	1000000000000
K_Q02bAU			3	Disability Support P	0100000000000
K_Q02bAU			4	Newstart Allowance	0010000000000
K_Q02bAU			5	Carer Payment	0001000000000
K_Q02bAU			6	Partner Allowance	0000100000000
K_Q02bAU			7	Widow Allowance from	0000010000000
K_Q02bAU			8	Wife Pension	0000001000000
K_Q02bAU			9	Mature Age Allowance	0000000100000
K_Q02bAU			10	Sickness Allowance	0000000010000
K_Q02bAU			11	Special Benefit	0000000001000
K_Q02bAU			12	No/None of these	0000000000100
K_Q02bAU			96	Valid skip	0000000000010
K_Q02bAU2	8	Income - Current pension - Period	-1	Missing	0000001
K_Q02bAU2			1	Week	0000000
K_Q02bAU2			2	Fortnight	1000000
K_Q02bAU2			3	Four weeks	0100000
K_Q02bAU2			4	Calendar month	0010000
K_Q02bAU2			5	Year	0001000
K_Q02bAU2			6	Other (please specif	0000100
K_Q02bAU2			96	Valid skip	0000010
K_Q03AU_01	4	Income - Current pensions2	-1	Missing	001
K_Q03AU_01			1	Marked	000
K_Q03AU_01			2	Not marked	100
K_Q03AU_01			6	Valid skip	010
K_Q03AU_02	4	Income - Current pensions2	-1	Missing	001
K_Q03AU_02			1	Marked	000
K_Q03AU_02			2	Not marked	100
K_Q03AU_02			6	Valid skip	010
K_Q03AU_03	4	Income - Current pensions2	-1	Missing	001
K_Q03AU_03			1	Marked	000
K_Q03AU_03			2	Not marked	100
K_Q03AU_03			6	Valid skip	010
K_Q03AU_04	4	Income - Current pensions2	-1	Missing	001
K_Q03AU_04			1	Marked	000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
K_Q03AU_04			2	Not marked	100
K_Q03AU_04			6	Valid skip	010
K_Q03AU_05	4	Income - Current pensions2	-1	Missing	001
K_Q03AU_05			1	Marked	000
K_Q03AU_05			2	Not marked	100
K_Q03AU_05			6	Valid skip	010
K_Q03AU_06	4	Income - Current pensions2	-1	Missing	001
K_Q03AU_06			1	Marked	000
K_Q03AU_06			2	Not marked	100
K_Q03AU_06			6	Valid skip	010
K_Q03AU_07	4	Income - Current pensions2	-1	Missing	001
K_Q03AU_07			1	Marked	000
K_Q03AU_07			2	Not marked	100
K_Q03AU_07			6	Valid skip	010
K_Q03AU_08	4	Income - Current pensions2	-1	Missing	001
K_Q03AU_08			1	Marked	000
K_Q03AU_08			2	Not marked	100
K_Q03AU_08			6	Valid skip	010
K_Q03AU_09	4	Income - Current pensions2	-1	Missing	001
K_Q03AU_09			1	Marked	000
K_Q03AU_09			2	Not marked	100
K_Q03AU_09			6	Valid skip	010
K_Q03cAU	8	Income - Current pension2 - Period	-1	Missing	0000001
K_Q03cAU			1	Week	0000000
K_Q03cAU			2	Fortnight	1000000
K_Q03cAU			3	Four weeks	0100000
K_Q03cAU			4	Calendar month	0010000
K_Q03cAU			5	Year	0001000
K_Q03cAU			6	Other (please specif	0000100
K_Q03cAU			96	Valid skip	0000010
K_Q03eAU	8	Income - Family Tax Benefit - Period	-1	Missing	0000001
K_Q03eAU			1	Week	0000000
K_Q03eAU			2	Fortnight	1000000
K_Q03eAU			3	Four weeks	0100000
K_Q03eAU			4	Calendar month	0010000
K_Q03eAU			5	Year	0001000
K_Q03eAU			6	Other (please specif	0000100
K_Q03eAU			96	Valid skip	0000010
K_Q04aAU2	8	Income - Child Support or Maintenance - Period	-1	Missing	0000001
K_Q04aAU2			1	Week	0000000
K_Q04aAU2			2	Fortnight	1000000
K_Q04aAU2			3	Four weeks	0100000
K_Q04aAU2			4	Calendar month	0010000
K_Q04aAU2			5	Year	0001000
K_Q04aAU2			6	Other (please specif	0000100

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
K_Q04aAU2			96	Valid skip	0000010
K_Q04AU_01	4	Income - Other listed sources	-1	Missing	001
K_Q04AU_01			1	Marked	000
K_Q04AU_01			2	Not marked	100
K_Q04AU_01			6	Valid skip	010
K_Q04AU_02	4	Income - Other listed sources	-1	Missing	001
K_Q04AU_02			1	Marked	000
K_Q04AU_02			2	Not marked	100
K_Q04AU_02			6	Valid skip	010
K_Q04AU_03	4	Income - Other listed sources	-1	Missing	001
K_Q04AU_03			1	Marked	000
K_Q04AU_03			2	Not marked	100
K_Q04AU_03			6	Valid skip	010
K_Q04AU_04	4	Income - Other listed sources	-1	Missing	001
K_Q04AU_04			1	Marked	000
K_Q04AU_04			2	Not marked	100
K_Q04AU_04			6	Valid skip	010
K_Q04bAU2	8	Income - Superannuation, annuity or private pensio	-1	Missing	0000001
K_Q04bAU2			1	Week	0000000
K_Q04bAU2			2	Fortnight	1000000
K_Q04bAU2			3	Four weeks	0100000
K_Q04bAU2			4	Calendar month	0010000
K_Q04bAU2			5	Year	0001000
K_Q04bAU2			6	Other (please specif	0000100
K_Q04bAU2			96	Valid skip	0000010
K_Q04cAU2	8	Income - Workers' compensation - Period	-1	Missing	0000001
K_Q04cAU2			1	Week	0000000
K_Q04cAU2			2	Fortnight	1000000
K_Q04cAU2			3	Four weeks	0100000
K_Q04cAU2			4	Calendar month	0010000
K_Q04cAU2			5	Year	0001000
K_Q04cAU2			6	Other (please specif	0000100
K_Q04cAU2			96	Valid skip	0000010
K_Q05aAU	5	Income - Rental investment property - profit/loss	-1	Missing	0001
K_Q05aAU			1	Profit	0000
K_Q05aAU			2	Loss	1000
K_Q05aAU			3	Neither	0100
K_Q05aAU			6	Valid skip	0010
K_Q05AU	4	Income - Rental investment property	-1	Missing	001
K_Q05AU			1	Yes	000
K_Q05AU			2	No	100
K_Q05AU			6	Valid skip	010
K_Q06AU1	6	Income - Business - current fin year - profit/loss	-1	Missing	00001
K_Q06AU1			1	Profit	00000
K_Q06AU1			2	Loss	10000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
K_Q06AU1			3	Neither	01000
K_Q06AU1			4	Previously reported	00100
K_Q06AU1			6	Valid skip	00010
K_Q06AU4	4	Income - Other business	-1	Missing	001
K_Q06AU4			1	Yes	000
K_Q06AU4			2	No	100
K_Q06AU4			6	Valid skip	010
K_Q06AU5	5	Income - Other business - profit/loss	-1	Missing	0001
K_Q06AU5			1	Profit	0000
K_Q06AU5			2	Loss	1000
K_Q06AU5			3	Neither	0100
K_Q06AU5			6	Valid skip	0010
K_Q07AU	4	Income - Shares	-1	Missing	001
K_Q07AU			1	Yes	000
K_Q07AU			2	No	100
K_Q07AU			6	Valid skip	010
K_Q07bAU	4	Income - Shares - less than $100	-1	Missing	001
K_Q07bAU			1	Yes	000
K_Q07bAU			2	No	100
K_Q07bAU			6	Valid skip	010
K_Q08AU	4	Income - Interest	-1	Missing	001
K_Q08AU			1	Yes	000
K_Q08AU			2	No	100
K_Q08AU			6	Valid skip	010
K_Q08bAU	4	Income - Interest - less than $100	-1	Missing	001
K_Q08bAU			1	Yes	000
K_Q08bAU			2	No	100
K_Q08bAU			6	Valid skip	010
K_Q09AU	4	Income - Any other sources	-1	Missing	001
K_Q09AU			1	Yes	000
K_Q09AU			2	No	100
K_Q09AU			6	Valid skip	010
K_Q09bAU	8	Income - Any other sources - Period	-1	Missing	0000001
K_Q09bAU			1	Week	0000000
K_Q09bAU			2	Fortnight	1000000
K_Q09bAU			3	Four weeks	0100000
K_Q09bAU			4	Calendar month	0010000
K_Q09bAU			5	Year	0001000
K_Q09bAU			6	Other (please specif	0000100
K_Q09bAU			96	Valid skip	0000010
K_Q11AU	11	Income - Main source	-1	Missing	0000000001
K_Q11AU			1	Wages or salary, inc	0000000000
K_Q11AU			2	Government pension o	1000000000
K_Q11AU			3	Child support or mai	0100000000
K_Q11AU			4	Superannuation, an a	0010000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
K_Q11AU			5	Workers' compensatio	0001000000
K_Q11AU			6	Profit or loss from	0000100000
K_Q11AU			7	Profit or loss from	0000010000
K_Q11AU			8	Dividends from share	0000001000
K_Q11AU			9	Other	0000000100
K_Q11AU			96	Valid skip	0000000010
K_Q13AU	4	Income - Other members of household	-1	Missing	001
K_Q13AU			1	Amount	000
K_Q13AU			2	Nil	100
K_Q13AU			6	Valid skip	010
K_Q13bAU	4	Income - Other members of household - profit/loss	-1	Missing	001
K_Q13bAU			1	Profit	000
K_Q13bAU			2	Loss	100
K_Q13bAU			6	Valid skip	010
K_Q13cAU	8	Income - Other members of household - Period	-1	Missing	0000001
K_Q13cAU			1	Week	0000000
K_Q13cAU			2	Fortnight	1000000
K_Q13cAU			3	Four weeks	0100000
K_Q13cAU			4	Calendar month	0010000
K_Q13cAU			5	Year	0001000
K_Q13cAU			6	Other (please specif	0000100
K_Q13cAU			96	Valid skip	0000010
K_Q14AU	11	Income - Other members of household - Main source	-1	Missing	0000000001
K_Q14AU			1	Wages or salary, inc	0000000000
K_Q14AU			2	Government pension o	1000000000
K_Q14AU			3	Child support or mai	0100000000
K_Q14AU			4	Superannuation, an a	0010000000
K_Q14AU			5	Workers' compensatio	0001000000
K_Q14AU			6	Profit or loss from	0000100000
K_Q14AU			7	Profit or loss from	0000010000
K_Q14AU			8	Dividends from share	0000001000
K_Q14AU			9	Other	0000000100
K_Q14AU			96	Valid skip	0000000010
NumUR15overAU	17	Number of usual residents 15 and over	-1	Missing	0000000000000001
NumUR15overAU			1	1	0000000000000000
NumUR15overAU			2	2	1000000000000000
NumUR15overAU			3	3	0100000000000000
NumUR15overAU			4	4	0010000000000000
NumUR15overAU			5	5	0001000000000000
NumUR15overAU			6	6	0000100000000000
NumUR15overAU			7	7	0000010000000000
NumUR15overAU			8	8	0000001000000000
NumUR15overAU			9	9	0000000100000000
NumUR15overAU			10	10	0000000010000000
NumUR15overAU			11	11	0000000001000000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
NumUR15overAU			12	12	0000000000100000
NumUR15overAU			13	13	0000000000010000
NumUR15overAU			14	14	0000000000001000
NumUR15overAU			15	15	0000000000000100
NumUR15overAU			96	Valid skip	0000000000000010
NumURAU	17	Number of usual residents	-1	Missing	0000000000000001
NumURAU			1	1	0000000000000000
NumURAU			2	2	1000000000000000
NumURAU			3	3	0100000000000000
NumURAU			4	4	0010000000000000
NumURAU			5	5	0001000000000000
NumURAU			6	6	0000100000000000
NumURAU			7	7	0000010000000000
NumURAU			8	8	0000001000000000
NumURAU			9	9	0000000100000000
NumURAU			10	10	0000000010000000
NumURAU			11	11	0000000001000000
NumURAU			12	12	0000000000100000
NumURAU			13	13	0000000000010000
NumURAU			14	14	0000000000001000
NumURAU			15	15	0000000000000100
NumURAU			96	Valid skip	0000000000000010
PARENTAU	4	Is a parent	-1	Missing	001
PARENTAU			0	False	000
PARENTAU			1	True	100
PARENTAU			6	Valid skip	010
Prov	17	Respondent province - From CMS	-1	Missing	0000000000000001
Prov			10	Newfoundland	0000000000000000
Prov			11	Prince Edward Island	1000000000000000
Prov			12	Nova Scotia	0100000000000000
Prov			13	New Brunswick	0010000000000000
Prov			24	Quebec	0001000000000000
Prov			35	Ontario	0000100000000000
Prov			46	Manitoba	0000010000000000
Prov			47	Saskatchewan	0000001000000000
Prov			48	Alberta	0000000100000000
Prov			59	British Columbia	0000000010000000
Prov			60	Yukon	0000000001000000
Prov			61	Northwest Territorie	0000000000100000
Prov			62	Nunavut	0000000000010000
Prov			76	U.S.A.	0000000000001000
Prov			77	Outside Canada/U.S.A	0000000000000100
Prov			96	Valid skip	0000000000000010
RESPONDENTAU	4	Selected person completed screener	-1	Missing	001
RESPONDENTAU			1	Yes	000

ITEM_ID	N Contrast	LABEL	VALUE	Category Label	CONTRAST
RESPONDENTAU			2	No	100
RESPONDENTAU			6	Valid skip	010
SecSchAU	4	Attending secondary school	-1	Missing	001
SecSchAU			1	Yes	000
SecSchAU			2	No	100
SecSchAU			6	Valid skip	010
TerSchAU	4	Attending full-time tertiary study	-1	Missing	001
TerSchAU			1	Yes	000
TerSchAU			2	No	100
TerSchAU			6	Valid skip	010

Appendix 3: Design Effect Tables

	Literacy Scale overall		
CNTRYID	Average	(s.e.)	Design Effect
Australia	280.40	0.91	2.39
Austria	269.45	0.74	1.41
Canada	273.49	0.57	3.45
Cyprus*	268.84	0.75	1.54
Czech Republic	274.01	0.98	3.53
Denmark	270.79	0.62	1.24
England (UK)	272.58	1.05	2.33
England/N. Ireland (UK)	272.46	1.02	3.81
Estonia	275.88	0.72	2.00
Finland	287.55	0.67	0.94
Flanders (Belgium)	275.48	0.83	1.55
France	262.14	0.59	1.01
Germany	269.81	0.92	2.01
Ireland	266.54	0.92	2.25
Italy	250.48	1.09	2.75
Japan	296.24	0.68	1.54
Korea	272.56	0.58	1.31
Netherlands	284.01	0.71	1.10
Northern Ireland (UK)	268.70	1.93	6.62
Norway	278.43	0.61	0.83
Poland	266.90	0.60	1.48
Russian Federation*	275.23	2.73	15.77
Slovak Republic	273.85	0.62	1.35
Spain	251.79	0.71	1.27
Sweden	279.23	0.68	0.80
United States	269.81	1.05	2.21

* Please refer to notes A and B regarding Cyprus, and the note regarding the Russian Federation in the *Note to Readers* section of this report.

	Literacy Scale by Gender					
	Female			Male		
CNTRYID	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect
Australia	279.48	1.11	1.92	281.32	1.28	2.19
Austria	267.39	0.93	1.19	271.53	1.04	1.36
Canada	272.34	0.79	3.65	274.63	0.86	3.59
Cyprus*	269.60	0.97	1.57	267.99	1.18	1.46
Czech Republic	272.32	1.30	3.37	275.68	1.26	2.61
Denmark	271.00	0.80	1.12	270.58	1.03	1.55
England (UK)	271.21	1.33	2.25	273.96	1.41	1.71
England/N. Ireland (UK)	271.03	1.29	3.70	273.90	1.37	2.77
Estonia	276.64	0.81	1.45	275.06	1.09	1.96
Finland	289.15	0.99	1.09	285.96	1.21	1.49
Flanders (Belgium)	272.81	1.08	1.39	278.09	0.97	1.01
France	262.23	0.69	0.71	262.05	0.87	1.05
Germany	267.21	1.19	1.76	272.35	1.17	1.58
Ireland	265.43	1.10	1.91	267.71	1.17	1.55
Italy	250.61	1.32	2.25	250.36	1.50	2.32
Japan	294.69	1.01	1.80	297.78	0.88	1.18
Korea	269.43	0.87	1.57	275.72	0.75	1.02
Netherlands	280.92	0.94	1.00	287.06	1.08	1.22
Northern Ireland (UK)	265.62	1.83	3.80	271.89	2.39	3.88
Norway	276.43	0.91	0.92	280.34	0.97	1.05
Poland	270.08	0.86	1.61	263.66	0.97	1.81
Russian Federation*	277.37	2.88	12.42	272.90	2.98	6.03
Slovak Republic	274.22	0.82	1.30	273.47	0.86	1.21
Spain	249.45	1.04	1.42	254.11	1.00	1.17
Sweden	277.54	1.10	1.02	280.88	1.08	1.06
United States	269.47	1.33	2.02	270.16	1.21	1.31

* Please refer to notes A and B regarding Cyprus, and the note regarding the Russian Federation in the *Note to Readers* section of this report.

PIAAC Design Effects

	Literacy Scale by Age Group														
	24 or less			25-34			35-44			45-54			55 plus		
CNTRYID	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect
Australia	284.13	2.21	2.01	287.49	1.67	1.70	288.73	1.46	1.62	276.86	1.76	1.79	262.75	1.72	1.75
Austria	277.72	1.47	1.04	279.80	1.46	1.03	274.64	1.69	1.56	266.16	1.37	1.27	249.81	1.59	1.47
Canada	275.73	1.27	3.67	285.14	1.26	3.18	279.65	1.36	3.96	267.98	1.29	3.79	260.38	1.09	2.79
Cyprus*	267.14	1.67	1.25	275.13	1.72	1.84	269.92	1.55	1.42	270.03	1.66	1.49	260.67	1.61	1.38
Czech Republic	280.53	2.11	4.23	286.72	1.82	2.79	275.15	2.02	2.73	265.76	1.71	1.53	262.38	1.98	3.40
Denmark	276.06	1.32	1.07	282.06	1.75	1.17	281.11	1.65	1.69	265.50	1.41	1.28	252.42	1.05	1.37
England (UK)	265.45	2.37	1.76	280.10	2.13	1.93	279.19	1.62	1.23	271.25	1.82	1.49	265.33	1.98	1.91
England/N. Ireland (UK)	265.69	2.28	2.91	280.02	2.07	3.18	279.02	1.57	2.06	270.98	1.75	2.40	265.03	1.94	3.04
Estonia	287.07	1.28	1.35	285.90	1.66	2.01	277.75	1.21	1.21	268.79	1.42	1.58	260.62	1.51	2.07
Finland	296.71	1.86	1.66	308.87	1.73	1.42	298.78	2.07	1.70	283.62	1.81	1.41	259.73	1.45	1.42
Flanders (Belgium)	285.03	1.64	1.45	290.77	1.78	1.35	282.38	1.60	1.15	271.89	1.61	1.40	255.04	1.55	1.21
France	275.03	1.29	1.00	278.00	1.43	1.14	266.80	1.32	1.02	253.71	1.16	0.91	241.81	1.25	1.12
Germany	278.91	1.61	1.39	281.31	1.78	1.39	275.26	1.61	1.25	263.64	1.65	1.54	253.62	1.66	1.40
Ireland	270.57	1.82	1.48	275.62	1.51	1.54	271.09	1.75	2.05	259.30	2.09	2.09	250.51	1.81	1.65
Italy	260.80	2.72	2.06	260.24	2.21	1.86	252.77	1.91	2.37	248.78	1.82	1.87	233.36	2.21	2.71
Japan	299.42	1.56	1.52	309.21	1.74	2.23	307.01	1.01	1.02	297.06	1.50	1.57	273.35	1.60	1.89
Korea	292.94	1.72	2.84	289.53	1.16	1.31	277.55	1.20	1.67	258.60	1.35	1.76	244.10	1.43	1.39
Netherlands	294.61	1.64	1.41	298.07	2.00	1.42	293.98	1.84	1.63	277.24	1.74	1.49	260.80	1.57	1.37
Northern Ireland (UK)	272.35	2.72	2.10	277.62	2.87	2.88	273.92	2.33	2.45	262.49	2.63	2.71	255.11	3.22	3.69
Norway	275.04	1.43	1.05	288.53	1.85	1.22	288.16	1.56	1.12	277.45	1.52	1.26	261.87	1.47	1.11
Poland	281.48	1.07	2.95	277.19	1.49	2.07	268.11	1.91	1.42	259.09	1.69	1.07	249.12	1.72	1.40
Russian Federation*	274.03	3.98	11.08	272.79	4.06	6.96	277.70	3.86	4.79	277.22	3.69	4.44	274.73	3.88	5.51
Slovak Republic	276.00	1.61	1.88	278.36	1.45	1.49	278.32	1.37	1.26	270.08	1.30	1.26	265.97	1.27	1.23
Spain	263.88	1.57	1.37	262.80	1.48	1.23	259.57	1.33	1.15	248.48	1.53	1.26	226.73	1.87	1.55
Sweden	282.76	1.68	1.13	290.01	1.92	0.99	287.39	1.85	1.06	276.01	1.68	1.07	262.37	1.33	0.91
United States	271.53	2.00	1.78	275.48	1.96	1.52	273.38	1.83	1.31	265.93	1.69	1.16	262.89	1.54	1.00

* Please refer to notes A and B regarding Cyprus, and the note regarding the Russian Federation in the *Note to Readers* section of this report.

	Numeracy Scale Overall		
CNTRYID	Average	(s.e.)	Design Effect
Australia	267.63	0.95	2.06
Austria	275.04	0.88	1.61
Canada	265.46	0.71	4.39
Cyprus*	264.63	0.79	1.25
Czech Republic	275.73	0.93	2.75
Denmark	278.28	0.73	1.47
England (UK)	261.81	1.10	2.03
England/N. Ireland (UK)	261.73	1.07	3.33
Estonia	273.12	0.53	1.02
Finland	282.23	0.70	1.00
Flanders (Belgium)	280.39	0.83	1.34
France	254.19	0.61	0.81
Germany	271.73	1.00	1.89
Ireland	255.59	1.02	2.16
Italy	247.13	1.06	2.08
Japan	288.17	0.74	1.48
Korea	263.39	0.69	1.52
Netherlands	280.35	0.71	0.99
Northern Ireland (UK)	259.17	1.82	4.71
Norway	278.30	0.79	1.05
Poland	259.77	0.82	2.47
Russian Federation*	269.93	2.74	16.62
Slovak Republic	275.81	0.79	1.58
Spain	245.82	0.62	0.88
Sweden	279.05	0.82	0.99
United States	252.84	1.17	2.05

* Please refer to notes A and B regarding Cyprus, and the note regarding the Russian Federation in the *Note to Readers* section of this report.

	Numeracy Scale by Gender					
	Female			Male		
CNTRYID	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect
Australia	260.77	1.20	1.84	274.47	1.42	2.14
Austria	268.47	1.14	1.48	281.66	1.20	1.42
Canada	258.17	0.95	4.46	272.75	0.90	3.23
Cyprus*	261.19	1.17	1.65	268.46	1.13	1.02
Czech Republic	271.19	1.30	2.95	280.20	1.36	2.71
Denmark	273.09	0.95	1.40	283.40	1.20	1.83
England (UK)	254.70	1.47	2.22	268.97	1.43	1.42
England/N. Ireland (UK)	254.62	1.42	3.68	268.88	1.39	2.29
Estonia	270.26	0.80	1.39	276.24	0.86	1.15
Finland	277.11	1.00	1.08	287.29	1.20	1.38
Flanders (Belgium)	272.28	1.15	1.42	288.31	1.14	1.21
France	248.92	0.89	0.90	259.72	0.88	0.82
Germany	262.99	1.32	1.75	280.28	1.31	1.66
Ireland	249.76	1.33	2.18	261.68	1.29	1.47
Italy	241.76	1.38	1.93	252.50	1.39	1.65
Japan	281.98	1.06	1.75	294.29	1.13	1.53
Korea	258.27	0.99	1.71	268.56	0.90	1.21
Netherlands	271.94	0.97	1.00	288.68	1.09	1.14
Northern Ireland (UK)	252.25	2.10	4.16	266.33	2.10	2.40
Norway	270.72	1.07	0.99	285.55	1.17	1.17
Poland	258.83	0.90	1.62	260.73	1.24	2.53
Russian Federation*	271.41	2.77	12.11	268.32	3.31	7.65
Slovak Republic	274.62	0.97	1.30	277.00	1.08	1.35
Spain	239.54	0.95	1.13	252.04	0.99	1.04
Sweden	272.17	0.98	0.72	285.73	1.31	1.31
United States	245.96	1.46	1.85	260.05	1.27	1.08

* Please refer to notes A and B regarding Cyprus, and the note regarding the Russian Federation in the *Note to Readers* section of this report.

PIAAC Design Effects

	Numeracy Scale by Age Group														
	24 or less			25-34			35-44			45-54			55 plus		
CNTRYID	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect
Australia	270.06	2.55	2.08	275.05	1.82	1.64	275.85	1.69	1.64	264.66	1.83	1.51	250.43	2.00	1.92
Austria	279.27	1.63	1.10	282.06	1.73	1.10	281.35	2.01	1.72	274.48	1.67	1.49	257.48	1.74	1.26
Canada	268.33	1.55	4.22	276.50	1.43	3.43	271.87	1.47	3.87	260.69	1.41	3.82	251.40	1.41	3.82
Cyprus*	264.21	2.07	1.43	273.14	2.00	1.94	268.96	1.63	1.24	264.56	1.76	1.20	250.17	1.75	1.18
Czech Republic	277.99	1.64	2.21	288.37	1.77	2.15	277.36	1.75	1.83	271.88	2.25	2.35	263.21	1.95	2.80
Denmark	273.09	1.54	1.20	286.72	1.89	1.18	290.01	1.60	1.35	276.79	1.60	1.39	265.35	1.20	1.53
England (UK)	256.27	2.68	1.90	266.72	2.24	1.60	268.84	1.90	1.35	259.10	1.93	1.32	256.93	1.92	1.40
England/N. Ireland (UK)	256.53	2.60	3.20	266.75	2.18	2.64	268.74	1.85	2.30	258.87	1.87	2.16	256.58	1.87	2.21
Estonia	278.54	1.22	1.10	283.63	1.69	1.95	275.10	1.13	0.97	268.96	1.44	1.54	259.44	1.26	1.39
Finland	284.77	1.83	1.34	302.45	2.08	1.90	292.03	2.15	1.69	279.27	1.97	1.46	260.05	1.26	1.01
Flanders (Belgium)	282.82	1.74	1.33	295.01	1.86	1.34	289.32	1.78	1.22	280.34	1.87	1.57	259.87	1.59	1.09
France	263.36	1.55	1.14	269.36	1.45	0.88	262.07	1.57	1.05	245.99	1.42	0.98	234.13	1.47	1.17
Germany	275.10	1.81	1.48	281.97	1.78	1.26	278.62	2.01	1.51	268.21	1.93	1.57	256.38	1.91	1.31
Ireland	257.87	2.25	1.70	265.50	1.65	1.47	260.48	1.74	1.53	249.59	2.11	1.65	238.27	2.34	2.06
Italy	251.30	2.63	1.63	262.41	2.28	1.67	250.88	1.88	1.86	243.71	1.95	1.62	229.37	2.21	2.06
Japan	283.21	2.29	2.30	297.32	1.64	1.51	296.64	1.33	1.27	291.47	1.71	1.52	273.22	1.62	1.56
Korea	280.92	1.91	2.83	280.69	1.37	1.50	270.64	1.48	2.13	251.06	1.42	1.57	231.76	1.67	1.52
Netherlands	285.40	1.76	1.36	292.98	1.81	1.07	287.38	2.08	1.72	277.10	1.70	1.22	261.98	1.66	1.30
Northern Ireland (UK)	263.59	3.40	2.64	267.58	2.90	2.42	265.78	2.43	2.20	251.64	2.13	1.40	245.18	3.06	2.54
Norway	270.93	1.73	1.14	284.93	2.02	1.09	289.02	1.89	1.18	280.30	1.69	1.17	264.72	1.73	1.13
Poland	268.59	1.11	2.64	270.43	1.50	1.89	261.72	2.17	1.59	254.23	2.10	1.47	243.65	1.85	1.41
Russian Federation*	272.54	3.75	10.51	268.64	4.22	7.56	270.04	3.58	4.71	272.11	3.16	3.40	266.64	3.94	5.60
Slovak Republic	277.98	1.76	1.65	278.82	1.65	1.34	281.37	1.65	1.29	275.36	1.62	1.37	265.28	1.55	1.30
Spain	255.15	1.72	1.53	257.29	1.32	0.91	254.90	1.27	0.99	242.32	1.59	1.22	220.53	1.75	1.17
Sweden	278.21	1.73	1.01	287.75	1.95	0.88	286.11	2.04	1.10	276.31	2.28	1.54	268.26	1.69	1.20
United States	249.42	2.19	1.57	259.85	2.19	1.46	257.68	1.89	1.04	249.77	2.07	1.26	247.15	1.77	0.99

* Please refer to notes A and B regarding Cyprus, and the note regarding the Russian Federation in the *Note to Readers* section of this report.

	PSTRE Scale Overall		
CNTRYID	Average	(s.e.)	Design Effect
Australia	288.68	0.88	2.81
Austria	283.98	0.73	1.44
Canada	282.43	0.68	4.80
Czech Republic	282.99	1.10	2.87
Denmark	283.08	0.68	1.56
England (UK)	280.50	0.95	2.18
England/N. Ireland (UK)	280.33	0.93	3.59
Estonia	277.62	1.01	2.95
Finland	289.37	0.83	1.73
Flanders (Belgium)	280.76	0.82	1.45
Germany	282.58	1.04	2.58
Ireland	276.80	1.01	2.57
Japan	294.03	1.19	2.38
Korea	282.97	0.79	2.02
Netherlands	286.40	0.76	1.50
Northern Ireland (UK)	275.03	1.97	7.14
Norway	286.49	0.57	0.88
Poland	274.92	1.33	4.54
Russian Federation*	276.25	4.34	22.33
Slovak Republic	281.08	0.82	1.74
Sweden	287.77	0.65	0.86
United States	277.44	1.15	2.84

* Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

	PSTRE Scale by Gender					
	Female			Male		
CNTRYID	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect
Australia	288.61	1.18	2.77	288.74	1.21	2.36
Austria	279.22	1.04	1.45	288.56	0.95	1.25
Canada	281.47	0.98	5.60	283.39	0.77	2.74
Czech Republic	280.91	1.49	2.88	284.87	1.72	3.20
Denmark	280.92	0.95	1.68	285.30	0.96	1.44
England (UK)	275.87	1.09	1.76	285.09	1.44	2.05
England/N. Ireland (UK)	275.67	1.08	3.02	284.96	1.40	3.22
Estonia	275.64	1.24	2.50	279.81	1.34	2.28
Finland	287.62	1.01	1.34	291.13	1.14	1.56
Flanders (Belgium)	277.71	1.16	1.52	283.68	1.07	1.21
Germany	279.80	1.24	1.88	285.11	1.37	2.17
Ireland	274.15	1.16	1.99	279.70	1.40	2.15
Japan	289.41	1.56	2.09	297.83	1.40	1.66
Korea	279.98	1.18	2.43	285.87	0.91	1.24
Netherlands	282.21	0.96	1.22	290.44	1.14	1.72
Northern Ireland (UK)	269.42	2.32	6.57	280.78	2.12	3.13
Norway	283.37	0.87	1.01	289.45	0.86	1.00
Poland	271.28	1.73	4.01	278.66	1.79	3.99
Russian Federation*	279.26	3.83	12.30	273.08	5.67	12.48
Slovak Republic	280.27	1.01	1.39	281.88	1.31	2.09
Sweden	285.58	0.98	1.01	289.88	1.01	1.03
United States	275.08	1.36	2.36	279.99	1.42	1.84

* Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

	PSL Scale by Age Group														
	24 or less			25-34			35-44			45-54			55 plus		
CNTRYID	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect	Average	(s.e.)	Design Effect
Australia	295.46	2.15	2.55	295.52	1.59	2.03	291.15	1.36	1.73	283.26	1.90	2.83	269.97	1.84	2.36
Austria	294.22	1.42	1.25	296.37	1.55	1.55	284.65	1.64	1.72	274.49	1.47	1.44	259.56	1.81	1.36
Canada	293.84	1.42	5.24	292.01	1.54	4.97	287.54	1.43	4.78	273.78	1.28	3.78	261.19	1.40	4.00
Czech Republic	296.68	2.15	4.11	297.04	1.67	1.79	276.59	2.55	2.61	269.53	2.56	2.20	263.00	2.83	3.21
Denmark	293.55	1.40	1.42	302.79	1.50	1.21	290.72	1.27	1.18	274.67	1.58	1.90	254.37	1.42	2.63
England (UK)	287.79	1.93	1.75	292.07	1.81	1.83	283.26	1.50	1.33	271.89	1.81	1.54	263.00	2.02	1.97
England/N. Ireland (UK)	287.76	1.89	2.98	291.84	1.76	3.05	282.98	1.46	2.22	271.62	1.78	2.47	262.76	1.97	2.97
Estonia	293.30	1.57	2.16	288.89	1.55	1.82	274.60	1.31	1.21	259.42	1.77	1.70	249.35	1.90	1.65
Finland	302.90	1.92	2.79	310.22	1.82	2.25	296.41	1.66	1.65	277.43	1.52	1.40	253.07	1.64	1.69
Flanders (Belgium)	298.95	1.66	1.72	297.04	1.62	1.33	285.58	1.65	1.36	269.51	1.61	1.48	253.26	2.05	1.61
Germany	294.81	1.79	2.09	295.51	2.01	1.95	285.49	1.77	1.66	273.11	1.72	1.70	259.81	2.40	2.20
Ireland	285.68	1.75	1.59	284.73	1.63	2.01	274.67	1.64	1.88	266.33	2.14	1.98	251.49	2.33	1.63
Japan	299.94	2.12	1.60	309.67	1.95	1.80	301.74	1.71	1.54	282.47	2.34	1.91	261.88	3.02	2.05
Korea	303.53	1.48	2.27	292.94	1.57	2.39	276.66	1.30	1.83	261.49	1.82	1.99	255.69	2.82	2.04
Netherlands	300.08	1.77	2.01	300.68	1.93	1.66	292.61	1.66	1.75	277.53	1.62	1.74	260.81	1.69	1.80
Northern Ireland (UK)	287.20	2.94	3.23	285.04	2.50	2.87	274.51	2.43	2.92	262.50	2.82	2.93	253.47	2.90	2.68
Norway	295.66	1.39	1.35	301.63	1.53	1.34	292.58	1.24	1.03	277.30	1.35	1.18	259.21	1.76	1.54
Poland	286.78	1.35	3.49	280.23	2.29	3.34	271.28	3.12	1.72	257.63	3.29	1.44	244.12	4.13	1.80
Russian Federation*	282.84	5.07	14.76	277.30	7.10	13.99	268.34	5.03	4.31	283.87	3.65	2.25	258.86	7.19	4.63
Slovak Republic	286.84	1.61	2.21	284.50	1.70	1.91	279.14	2.07	2.15	274.78	2.36	2.29	271.24	2.46	1.64
Sweden	301.91	1.66	1.72	304.71	1.50	1.03	293.64	1.73	1.22	278.33	1.65	1.22	259.25	1.47	1.09
United States	285.16	2.24	2.55	283.41	2.03	1.95	278.97	2.22	2.00	270.70	1.73	1.31	266.77	2.53	2.69

* Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Appendix 4: PIAAC-IALS-ALL Trend Variables

Appendix 4: PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
Section A. General Information + Section B Education and Training + Section J Background Information	**Section A. General Information**	**Section A. General Information**		`
Date of birth				
☑A_Q01a. Can you please tell me in which year you were born? ☑A_Q01b And in which month were you born?		☑AA1. On what date were you born?	TREND	
Gender				
≈A_N01. Is the respondent male or female?		≈AA2. Is the respondent male or female?	TREND	
Respondent's origin				
≈J_Q04a. Were you born in #insert country name#?	≈Were you born in #insert country name#?	≈Were you born in #insert country name#?	TREND	
≈J_Q04b. In what country were you born?	≈A2. In what country were you born?	≈A1D. In what country were you born?	TREND	
☑J_Q04c. At what age or in which year did you first immigrate to #insert country name#?	≈A3. In what year did you first immigrate to #insert country name#?	≈A2. In what year did you first immigrate to #insert country name#?	TREND	

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables

☑ means exact same question;
≈means similar question but slightly different wording;
* means same concept but different wording/answer categories

PIAAC BQ MS version 2.1 d.d. 15-12-2010	IALS BQ	ALL BQ	Comparison status	Notes
Educational background - Formal education				
☑Derived variable on years of schooling	≈A7. During your lifetime, how many years of formal education have you completed beginning with grade one and not counting repeated years at the same level?	≈A3. During your lifetime, how many years of formal education have you completed beginning with grade one and not counting repeated years at the same level?	TREND	
☑B_Q01b Which of the qualifications on this card is the highest you have obtained?	☑A8. What is the highest level of schooling you have ever completed?	☑A4A. Have you graduated from high school? ☑A4B. What is the highest grade of elementary or high school that you have ever completed? ☑A4C. What is the highest level of schooling that you have ever completed?	TREND	Derived variable in three categories ISCED 1 and 2; ISCED 3 and 4 and ISCED 5 and 6
☑B_Q01c When you completed this qualification, how old were you, or what year was it?		☑A6. How old were you when you completed your <highest level of schooling completed>?	TREND	

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables

☑ means exact same question;
≈means similar question but slightly different wording;
* means same concept but different wording/answer categories

PIAAC BQ MS version 2.1 d.d. 15-12-2010	IALS BQ	ALL BQ	Comparison status	Notes
Section J Background Information	**Section B. Linguistic Information**	**Section B. Linguistic Information**		
Language background				
≈J_Q05a What is the language that you first learned at home in childhood AND STILL UNDERSTAND?	☑B8. What language did you first learn to read and write?	≈B1. What is the language that you first learned at home in childhood and still understand?	TREND	
≈J_Q05b What language do you speak most often at home?	≈B14. What language do you speak most often at home?	≈B2. What language do you speak most often at home?	TREND	
Section J Background Information	**Section C. Parental Information**	**Section C. Parental Information**		
Respondent's mother's background				
≈J_Q6a Was your mother or female guardian born in #insert country name#?	≈C2. Was your mother (female guardian) born in #insert country name#?	≈Was your mother or female guardian born in #insert country name#?	TREND	
≈J_Q06b What was the highest level of education your mother or female guardian ever completed?	≈C5. What was the highest level of schooling that your mother (female guardian) ever completed?	≈C2. What was the highest level of schooling that your mother or female guardian ever completed?	TREND	
Respondent's father's background				
≈J_Q7a Was your father or male guardian born in #insert country name#?	≈C8. Was your father (male guardian) born in #insert country name#?	≈Was your father or male guardian born in #insert country name#?	TREND	
≈J_Q07b What was the highest level of education your father or male guardian ever completed?	≈C11. What was the highest level of schooling that your father (male guardian) ever completed?	≈C6. What was the highest level of schooling that your father or male guardian ever completed?	TREND	

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables

☑ means exact same question;
≈means similar question but slightly different wording;
* means same concept but different wording/answer categories

PIAAC BQ MS version 2.1 d.d. 15-12-2010	IALS BQ	ALL BQ	Comparison status	Notes
Section C Current Status and Work History	**Section D. Labor Force Information**	**Section D. Labor Force Activities**		
Respondent's employment status				
*C_Q07 Please look at this card and tell me which ONE of the statements best describes your current situation. If more than one statement applies to you, please indicate the statement that best describes how you see yourself.	☑D1. I would now like to talk about your employment status. What is your current work situation? Are you employed, retired, unemployed / looking for work, a student (including Work Programs), homemaker or other?	☑D1. I would now like to talk about your employment status. What is your current work situation? Are you now employed or self employed, not working and looking for work, retired, a student (including work programs), doing unpaid household work or other?	TREND	Combine PIAAC(1,2) to ALL(1), PIAAC(4.5) to ALL(4), PIAAC(7,8,10) to ALL(6)

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
Work history - Past 12 months				
☑C_Q08b During the last 12 months, that is since ^MonthYear, did you have any paid work? Please include self-employment.	☑D2. Did you work at a job or business at any time in the past 12 months (regardless of the number of hours per week)?	☑D2. Did you work at a job or business at any time in the last 12 months; that is, from <month and year> to <month and year> (regardless of the number of hours per week)? PLEASE INCLUDE AS WORK TIME OFF FOR VACATION, ILLNESS, MATERNITY/PATERNITY LEAVE, STRIKES AND LOCKOUTS.	TREND	PIAAC CD09(1,2) to D2(1); CD09(3,4) to D2(2);CD09(5) to D2(8 or 9)
Job information - Current job or last (past 12 months) job held				
☑D_Q02a In what kind of business, industry or service do you work? Please give a full description. D_Q02b What does your firm or organization mainly make or do? Please give a full description.	≈D8. What kind of business, industry or service was this? (Give full description, e.g. fish canning plant, automobile manufacturing plant, municipal government)	≈D26. What kind of business, industry or service was/is this? (Give full description, e.g. fish canning plant, automobile manufacturing plant, municipal government .)	Not comparable	

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables

☑ means exact same question;
≈means similar question but slightly different wording;
* means same concept but different wording/answer categories

PIAAC BQ MS version 2.1 d.d. 15-12-2010	IALS BQ	ALL BQ	Comparison status	Notes
☑D_Q01 What is your job title? What are your most important responsibilities? Please give a full description.	☑D9. What kind of work were you doing at this job? (Give full description or occupational title, e.g. office clerk, machine operator, computer programmer)	☑D27. What kind of work were/are you doing at this job? (Give full description or occupational title, e.g. office clerk, machine operator, computer programmer.). D28. What were/are your most important activities or duties? (Give full description e.g. filing documents, drying vegetables, forest examiner.)	TREND	ISCO 1 digit to make it comparable
*D_Q12a Still talking about your current job: If applying today, what would be the usual qualifications, if any, that someone would need to GET this type of job?		*D28B. What level of education was required to do your main job?	Not comparable	
*D_Q04 In this job, are you working as an employee or are you self-employed? D_Q07a. Do you have employees working for you? Please include family members working paid or unpaid in the business. *D7b. How many employees do you employ? Would that be .... D_Q08a Do you manage or supervise other employees? *D_Q08b. How many people do you supervise or manage, directly or indirectly?	≈D11. What was your status at this job? Was it as an employee without supervisory responsibilities, an employee with limited supervisory or management responsibilities (5 persons or less), an employee with more extensive supervisory or management responsibilities (more than 5 persons), a self-employed without employees, a self employed with employees or a family worker (unpaid)?	≈D29. What was/is your status at this job or business? Were/are you...an employee without supervisory responsibilities, an employee with supervisory or management responsibilities for up to 5 persons, an employee with supervisory or management responsibilities for more than 5 persons, self-employed without employees, self employed with employees or unpaid family worker?	TREND	OK with new derived variables

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
*D_Q09 What kind of employment contract do you have? Is that …	*D12. What type of job was this? Was or is this job a permanent job or work contract of unlimited duration or a temporary job or work contract of limited duration?		Not comparable	
☑D_Q10 How many hours do you usually work per week in this job? Include any usual paid or unpaid overtime, but exclude lunch breaks or other breaks	☑D13. How many hours per week did you usually work at this job?	☑D37. On average, how many hours per week did/do you usually work at this job or business? (If it varies greatly ask for the average of the last 4 weeks of work)	TREND	
☑D_Q06a How many people work for your employer at the place where you work? Would that be …	*D10. In total, about how many persons are employed by this business at all locations in #insert country name#?	☑D30. About how many persons were/are employed at the location where you work(ed)? Would it be less than 20, 20 to 99, 100 to 499, 500 to 999 or 1000 and over?	Not comparable	
☑D_Q16a What is the easiest way for you to tell us your usual gross wage or salary for your current job? Would it be …		☑D39. What is the easiest way for you to tell us your usual wage or salary for this job? Would it be hourly, weekly, annually or on some other basis?	Not comparable	

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
*D_Q16b + derived variables What is your usual gross pay ^PerHourDayEtc? Please give as good an approximation as you can. By gross, we mean before deductions for tax, social security contributions, and the like. Please include any regular overtime pay, regular bonuses, tips and commissions. Don't include annual bonuses such as 13th month or holiday pay.		*D41. What was/is your (interviewer fill text as indicated in D39, e.g. hourly, weekly, etc.) wage or salary before taxes and all other deductions at this job? Including tips and commissions?	TREND	OK with derived variable of quintile of yearly earnings
*D_Q18a What were your total earnings last ^YearMonth from your current business after deducting all business expenses, but before deducting income taxes, social security contributions, and the like?		*D43. What was/is your annual personal net income before taxes and deductions from this business – that is, after all business expenses?	TREND	OK with derived variable of quintile of yearly earnings
Skills used at work				
Section G. Skill use literacy, numeracy and ICT at work	**Section E. Reading and Writing at Work and Looking for Work**	**Section E. Literacy and Numeracy Practices at Work**		
☑CURRENTLY WORKING OR HAD PAID WORK IN THE LAST 12 MONTHS	☑*Section E. is for respondents who are employed now or who worked in the past 12 months (regardless of the number of hours per week).*	☑*Section E. is for respondents who are currently employed or who worked in the last 12 months.*	*TREND*	

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
☑G_R01 The following questions are about reading activities that you ^UndertakeUndertook as part of your ^JobLastjob. Please only report reading that ^IsWas part of your ^JobLastjob, not reading you ^DoDid in your non-work time. Include any reading you might do on computer screens or other electronic displays.	☑E1. The following questions refer to the job at which you worked the most hours in the last 12 months.	☑E1. The next questions are about your reading, writing and mathematics activities at your main job – whether these activities are done on paper or on computer.	TREND	
*G_Q01a. In your ^JobLastjob, how often ^DoDid you usually …	≈How often (do/did) you read or use information from each of the following as part of your main job? Would you say every day, a few times a week, once a week, less than once a week, rarely or never?	≈How often do/did you read or use information from one of the following as part of your main job? Would you say at least once a week, less than once a week, rarely or never.	TREND	PIAAC and ALL/IALS categorical equivalence: PIAAC(1)=ALL(4);PIAAC(2,3)=ALL(2,3);PIAAC(4,5)=ALL(1); RF and DK are the same
≈G_Q01a read directions or instructions ≈G_Q01b read letters, memos or e-mails? *G_Q01c read articles in newspapers, magazines or newsletters? ≈G_Q01f read manuals or reference materials? ≈G_Q01g read bills, invoices, bank statements or other financial statements? *G_Q01h read diagrams, maps or schematics?	≈A. Letters or memos *B. Reports, articles, magazines or journals ≈C. Manuals or reference books, including catalogues ≈D. Diagrams or schematics ≈E. Bills, invoices, spreadsheets or budget tables	≈a) Letters, memos or e-mails *b) Reports, articles, magazines, or journals ≈c) Manuals or reference books including catalogues *d) Diagrams or schematics ≈e) Directions or instructions ≈f) Bills, invoices, spreadsheets or budget tables		

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
☑The following questions are about writing activities that you ^UndertakeUndertook as part of your ^JobLastjob. Include any writing you might do on computers or other electronic devices.	☑E2. How often (do/did) you write or fill out each of the following as part of your main job? Would you say every day, a few times a week, once a week, less than once a week, rarely or never?	☑E2. How often do/did you write or fill out each of the following as part of your main job? Would you say at least once a week, less than once a week, rarely or never.	Not comparable	
☑In your ^JobLastjob, how often ^DoDid you usually … ≈G_Q02a. Write letters, memos or e-mails? ☑G_Q02b Write articles for newspapers, magazines or newsletters? ☑G_Q02c. Write reports? ☑G_Q02d. Fill in forms?	A. Letters or memos ☑B. Forms or things such as bills, invoices, or budgets ☑C. Reports or articles	a) Letters, memos or e-mails ☑b) Reports, articles, magazines or journals ☑e) Bills, invoices, spreadsheets or budget tables		
☑The following questions are about activities that you ^UndertakeUndertook as part of your ^JobLastjob and that involve numbers, quantities, numerical information, statistics or mathematics.	☑E3. In your main job, how often do you use arithmetic or mathematics (that is, adding, subtracting, multiplying or dividing) to:	☑E3. How often do/did you do each of the following as part of your main job? Would you say at least once a week, less than once a week, rarely or never.	Not comparable	
☑In your ^JobLastjob, how often ^DoDid you usually … ≈G_Q03b calculate prices, costs or budgets?	≈B) Calculate prices, costs or budgets?	≈b) Calculate prices, costs, or budgets		

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
Section B Education and Training	**Section F. Adult Education**	**Section F. Participation in Education and Learning**		
Education or training which the respondent has taken in the past 12 months				
☑Derived variable based on B_Q02a, B_D03d, B_Q04a and B_Q12	≈During the past 12 months, that is, since August 1993, did you receive any training or education including courses, private lessons, correspondence courses, workshops, on- the-job training, apprenticeship training, arts, crafts, recreation courses or any other training or education?	≈F1. During this time, did you take any education or training? This education or training would include programs, courses, private lessons, correspondence courses, workshops, on-the-job training, apprenticeship training, arts, crafts, recreation courses, or any other training or education?	TREND	if any B_Q02a or B_Q04a or B_Q12a, or B_Q12c or B_Q12e or B_Q12g is yes or if B_D03d=<12, then equivalent to ALL F1(yes).
☑Derived variable based on B_Q04b and B_Q12	☑F2. In total, how many courses did you take in the past 12 months?		Not comparable	
		**(Compared to IALS BQ, items are divided into "a program of studies" and "courses not a part of a program of studies".)*		

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
☑B_Q02a Are you currently studying for any kind of formal qualification? B_Q04a During the last 12 months, that is since ^MonthYear, have you studied for any formal qualification, either full-time or part-time?	☑F5. Were you taking this training or education towards a university degree/diploma/certificate, a college diploma/certificate, a trade-vocational diploma/certificate, an apprenticeship certificate, an elementary or secondary school diploma or professional or career upgrading?	☑F2. During the last 12 months, that is, from < month and year > to < month and year > did you take any courses as part of a PROGRAM of studies toward a certificate, diploma or degree? Examples would include a high school diploma; a trade/vocational diploma or registered apprenticeship certificate; a college or CEGEP diploma; a diploma granted from a program of studies at a private school; a university certificate, diploma or degree?	TREND	If any B_Q02a or B_Q04a is yes or if B_D03d=<12 then equivalent to ALL F2(yes)

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
☑B_Q05c Were the main reasons for choosing to study for this qualification job related?		☑F13. What was the main reason you took this program of studies? Was it for job or career-related reasons or personal interest such as hobby/leisure, volunteer activities, to improve some general skills (reading, writing) or for general education?	TREND	
*Derived variable based on B_Q12		*F15. During the last 12 months, did you participate in any courses that were NOT PART OF YOUR PROGRAM OF STUDIES?	TREND	If any B_Q12a, or B_Q12c or B_Q12e or B_Q12g is yes
Education or training wanted but not taken in the past 12 months				
☑B26a. In the last 12 months, were there ^MoreAny learning activities you wanted to participate in but did not? Include both learning activities that lead to formal qualifications and other organized learning activities.	≈F15. Since August 1993, was there any training or education that you WANTED to take for career or job-related reasons but did not?	≈F27. During the last 12 months, was there any training or education that you wanted to take for career or job-related reasons but did not?	TREND	if yes to either of 2 ALL/IALS questions then PIAAC B26a is yes, if no to both for IALS/ALL questions then PIAAC B26a is no

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
	≈F17. Since August 1993, was there any other training that you WANTED to take but did not, such as hobby, recreational or interest courses?	≈F28. During the last 12 months, was there any other training or education that you wanted to take but did not, such as hobby, recreational, or personal interest courses?		
Section H. Skill use literacy, numeracy and ICT in everyday life + Section I About yourself	**Section G. Reading and Writing General**	**Section G. Literacy and Numeracy Practices Generally, Social Capital and Well Being**		
Reading and writing in respondents' daily life				
☑H_R01 ^TalkedAboutWork I would now like to talk about your reading activities ^EverydayReading Include any reading you might do on computer screens or other electronic displays.	≈G7. How often (do/did) you read or use information from each of the following as part of your daily life? Would you say every day, a few times a week, once a week, less than once a week, rarely or never?	≈G3. How often do you read or use information from each of the following sources as part of your daily life? Please don't include time spent as part of your job or schooling. Would you say at least once a week, less than once a week, rarely or never?	TREND	PIAAC(1)=ALL(4);PIAAC(2,3)=ALL(2,3);PIAAC(4,5)=ALL(1); RF and DK are the same; Take maximum frequency of either G3a and G3b to make equivalent derived variable to PIAAC H_Q01c.
☑H_Q01 ^Ineverydaylife, how often do you usually …	☑A. Letters or memos.	≈a) How often do you read or use information from newspapers as part of your daily life? G3A2. How often do you read newspapers in < insert language >?		

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
☑H_Q01a. read directions or instructions?	☑B. Reports, articles, magazines or journals.	≈b) How often do you use or read information from magazines or articles as part of your daily life?		
≈H_Q01b. read letters, memos or e-mails?	☑C. Manuals or reference books, including catalogues.	≈c) How often do you read or use information from books –fiction or non-fiction as part of your daily life?		
≈H_Q01c. read articles in newspapers, magazines or newsletters?	☑D. Diagrams or schematics.	d) How often do you read or use information from letters, notes, e-mails as part of your daily life?		
≈H_Q01e read books, fiction or non-fiction?	☑E. Bills, invoices, spreadsheets or budget tables.			
☑H_Q01f read manuals or reference materials?	☑G. Directions or instructions for medicines, recipes, or other products.			
☑H_Q01g read bills, invoices, bank statements or other financial statements?				
☑H_Q01h read diagrams, maps, or schematics?				

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
Civic participation - volunteer work				
*Q05f. In the last 12 months, how often, if at all, did you do voluntary work, including unpaid work for a charity, political party, trade union or other non-profit organization?		*G8. The next questions are about your volunteer work and the organizations in which you participate. During the last 12 months did you participate in any of the following groups or organizations? *a) A political organization *e) A neighborhood, civic or community association or a school group <e.g. Parent/Teachers Association, your neighborhood community association>	Not comparable	
Health				
≈I_Q08. In general, would you say your health is excellent, very good, good, fair or poor?		≈G11. In general, would you say your health is?	TREND	

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables

☑ means exact same question;
≈means similar question but slightly different wording;
* means same concept but different wording/answer categories

PIAAC BQ MS version 2.1 d.d. 15-12-2010	IALS BQ	ALL BQ	Comparison status	Notes
Section G. Skill use literacy, numeracy and ICT at work + Section H. Skill use literacy, numeracy and ICT in everyday life		**Section H. Information and Communication Technology Literacy (ICTL)**		
Use of information technologies – computer use				
☑derived variable based on G_Q04a ^DoiDid you use a computer in your ^JobLastjob? H_Q04a Have you ever used a computer?		☑H2. Have you ever used a computer? H7. In the last 12 months, did you use a computer in your job? (If you have more than one job, tell us about the one at which you work the most hours)	TREND	YES only for the general question on computer experience. Use derived variable of computer experience of PIAAC
Section J. Background Information	**Section H. Family Literacy**	**Section K. Household Information and Income**		
	Section J. Household Information			
Respondents' children's education				
☑J_Q01. Including yourself, how many people usually live in your household? Please include people who are temporarily living elsewhere.''	≈J4. Including yourself, how many people live in this household?	≈K1. The next questions ask for general household information. Including yourself, how many people live in your household?	TREND	

Appendix 4 (cont.): PIAAC-IALS-ALL trend variables ☑ means exact same question; ≈means similar question but slightly different wording; * means same concept but different wording/answer categories				
PIAAC BQ MS version 2.1 d.d. 15-12-2010	**IALS BQ**	**ALL BQ**	**Comparison status**	**Notes**
*J_Q03b Do you have children? Please include stepchildren and children not living in your household.	*H1. Are you the parent or guardian of any children aged 6 to 18 that are presently living with you?	*K2. Do you have any DEPENDENT children living with you in your household? (Children for whom you are financially responsible and/or have sole or joint custody).	Not comparable	
*J_Q03c. How old is this child? *J_Q03d1. How old is your youngest child?	*H2. What is the age of your youngest child between 6 to 18 years of age?	*K3. What is the age of the youngest child in your household?	Not comparable	

Appendix 5: Mapping of ISCED Levels to Years of Schooling

Variable name	Variable: VET
Description	Label: Highest level of education attained at ISCED 3 or ISCED 4 level has vocational orientation.
Rationale	Rationale: For analysis of effects of education tracking it is useful to have an indicator of whether the highest level of education attained was in vocational or general education
Alert	Caveat: Users of the data must be aware that a scheme attaching the vocational orientation to the highest education degree obtained can be indicative only and neglects country differences and different traditions with regard to vocational education.
F	Derivation: Based on answers to national version of B_Q01a, supplemented by information provided by National Project Managers (NPMs) and OECD LSO network experts
Labels	Categories: VET=1 if highest level of education attained (only ISCED3 or 4) has vocational orientation; VET=0 if highest level of education attained (only ISCED3 or 4) has general/academic orientation; VET=.v distinction of orientation not applicable for this ISCED level, VET=.n information of orientation for all respondents in this country is missing because either PIAAC categories make ex-post distinction impossible or orientation information is missing for this country.

Variable name	Variable: YRSQUAL
Description	Label: Years of schooling associated with the highest level of education attained
Rationale	Rationale: For returns to education analyses it is useful to have an estimate of the years of schooling associated with the highest level of education attained
Prerequisites/input	Derivation: based on the answers to question B_Q01a, supplemented by information provided by National Project Managers (NPMs) and OECD LSO network experts on association with years of schooling and educational categories used to gather information on highest level of education attained.
Alert	Caveat: Users of the data must be aware that a scheme converting highest education degree obtained into years of schooling represents an oversimplification of the flexibility of national education systems (see also note of Australian Government Department of Education, Employment and Workplace Relations).
Labels	

Variable name	Variable: YRSGET
Description	Label: Years of schooling necessary to get current job
Rationale	Rationale: In combination with YRSQUAL, we can get an indication of vertical educational mismatch.

Prerequisites/input	Derivation: Based on the answers to question D_Q12s (qualification needed to get the job), supplemented by information provided by National Project Managers (NPMs) and OECD LSO network experts on association with years of schooling and educational categories used to gather information on level of education. Since years of schooling necessary to get the current job have been derived using the same information as for YRSQUAL, the difference between the two variables should provide a good indication of the respondent's vertical education-job match.
Alert	Caveat: Users of the data must be aware that a scheme converting qualification necessary to get the current job into years of schooling represents an oversimplification of the flexibility of national education systems (see also note of Australian Government Department of Education, Employment and Workplace Relations). The following countries did not use the same show cards (educational categories) in B_Q01a and D_Q12a: Canada, Germany and Sweden. In these countries, mismatch variables created by substracting YRSQUAL and YRSGET could be artificially non-zero. These countries should be excluded from any analysis using mismatch variables based on YRSQUAL and YRSGET.
Labels	

AUSTRALIA							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the following qualifications is the highest you have obtained?		`	What ^isWas the highest year of primary or secondary school you ^have completed?			5
	No formal qualification or below ISCED 1	**1**	No correspondence	No correspondence			
	ISCED 1	**2**	Primary school	Primary school	8		
	ISCED 2	**3**	Junior secondary school or Cert Certificate I, II	Junior secondary school or Cert Certificate I, II	12		
	ISCED 3C shorter than 2 years	**4**	Certificate III	Certificate III	14	V	
	ISCED 3C 2 years or more	**5**	No correspondence	No correspondence			
	ISCED 3A-B	**6**	Senior Secondary school	Senior Secondary school	14	A	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**	No correspondence	No correspondence			
	ISCED 4C	**8**	Certificate IV	Certificate IV	16	V	
	ISCED 4A-B	**9**	No correspondence	No correspondence			
	ISCED 4 (without distinction A-B-C)	**10**	No correspondence	No correspondence			
	ISCED 5B	**11**	Diploma, Advanced diploma and Associate Degree	Diploma, Advanced diploma and Associate Degree	16		
	ISCED 5A, bachelor degree	**12**	Bachelor Degree, Graduate Certificate and Graduate Diploma	Bachelor Degree, Graduate Certificate and Graduate Diploma	17		
	ISCED 5A, master degree	**13**	Master Degree level	Master Degree level	19		
	ISCED 6	**14**	Doctoral Degree level	Doctoral Degree level	22		

AUSTRIA							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the qualifications on this card is the highest you have obtained?		Was ist Ihre höchste abgeschlossene Schulbildung?	Which of the following qualifications is the highest you have obtained?			6
	No formal qualification or below ISCED 1	**1**		No correspondence			
	ISCED 1	**2**	Kein Pflichtschulabschluss	No compulsory school	7		
	ISCED 2	**3**	Pflichtschulabschluss	Compulsory school	8		
	ISCED 3C shorter than 2 years	**4**	Fach-oder Handelsschule: kürzer als 2 Jahre	Vocational School (< 2 Years)	9	V	
	ISCED 3C 2 years or more	**5**		No correspondence			
	ISCED 3A-B	**6**	Lehre mit Berufschule	Apprenticeship	12	V	
			Fach-oder Handelsschule: 2 Jahre und länger	Vocational School (2 Years and longer)	11	V	
			AHS (z.B. Gymnasium)	Academic Secondary School	12	A	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**		No correspondence			
	ISCED 4C	**8**					
	ISCED 4A-B	**9**	Fach-oder Handelsschule: Diplomkrankenpflege	Nursing	15	V	
			BHS (z.B. HAK, HTL, BAKIP)	Vocational college	13	V	
	ISCED 4 (without distinction A-B-C)	**10**		No correspondence			
	ISCED 5B	**11**	Meister- und Werkmeisterprüfung, Bauhandwerkerprüfung	Master craftsman's certificate	14		

AUSTRIA							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
			Kolleg, Abiturientenlehrgang	Post-secondary courses	14		
			Akademie (z.B. Pädak, SozAK, BPA, Med.-Tech. Akademie, LW, MilAK)	Post-secondary colleges	15		
			Universitäre Lehrgänge (ohne vorangeganges Studium)	**University courses**	**14**		
	ISCED 5A, bachelor degree	12	Universität oder Fachhochschule: Bakkalaureat/Bachelor	*University-Bachelor*	*15*		
	ISCED 5A, master degree	**13**	**Universität oder Fachhochschule: Magisterium/Master (Diplomstudium, Doktorat als Erstabschluss)**	University-Master	17		
			Postgraduale Universitätslehrgänge (z.B. MBA, MAS)	Post-graduate courses	19		
	ISCED 6	**14**	**Doktorat nach akademischem Erstabschluss**	Doctoral Programme	19		

CANADA						
International English Version	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
Which of the qualifications on this card is the highest you have obtained?						6
No formal qualification or below ISCED 1	**1**	Less than Grade 6				
		No formal education				
ISCED 1	**2**	Grade 6		6		
ISCED 2	**3**	Grade 7-8 (Secondary 1 or 2 in QUE)		9		
		Grade 9 (Secondary 3 in QUE or Senior 1 in MAN)		9		
		Grade 10 - 13 (Secondary 4 or 5 in QUE, Senior 2, 3 or 4 in MAN, Level I, II or III in NFLD, OAC in ONT)		9		
ISCED 3C shorter than 2 years	**4**	No correspondence				
ISCED 3C 2 years or more	**5**	No correspondence				
ISCED 3A-B	**6**	No correspondence				
ISCED 3 (without distinction A-B-C, 2y+)	**7**	High school diploma or equivalent		12	A	
ISCED 4C	**8**	Apprenticeship certificate		13	V	

CANADA						
International English Version	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
		Trade/vocational certificate (includes an attestation of vocational training, diploma of vocational studies or attestation of vocational specialization offered in Quebec) with duration of less than 2 years		12-13 (Quebec for range)	V	
		Non-university certificate or diploma from a college, school of nursing, technical institute with duration less than 2 years		12-13 (Quebec for range)	V	
		CEGEP diploma or certificate as part of university transfer program		13	V	
ISCED 4A-B	**9**	CEGEP diploma or certificate not part of a university transfer program with duration less than 2 years, only Quebec		13	A	
		University transfer program		14	A	
ISCED 4 (without distinction A-B-C)	10	No correspondence				

CANADA						
International English Version	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
ISCED 5B	11	CEGEP diploma or certificate not part of a university transfer program with duration 2 years or more, only Quebec		*14*		
		Trade/vocational certificate (includes an attestation of vocational training, diploma of vocational studies or attestation of vocational specialization offered in Quebec) with duration of 2 years or more, only outside Quebec		14		
		Non-university certificate or diploma from a college, school of nursing, technical institute with duration 2 years or more, only outside Quebec				
		University certificate or diploma below bachelor's degree		14		
ISCED 5A, bachelor degree	**12**	Bachelor's degree		16		
		University certificate above the bachelor's		16		

CANADA						
International English Version	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
		First professional degree (medical, veterinary medicine, dental, optometry, law, divinity).		16		
ISCED 5A, master degree	**13**	Master's		18		
ISCED 6	**14**	Ph.D.		22		

CYPRUS*							
Int. Question No	**International English Version**	**Int. Value Code**	**National version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the following qualifications is the highest you have obtained?			Which of the following qualifications is the highest you have obtained?			6
	No formal qualification or below ISCED 1	**1**	Δεν φοίτησα ποτέ σε σχολ	**I never went to school**			
	ISCED 1	**2**	Δημοτικό Σχολείο	**Primary school**	6		
	ISCED 2	**3**	Δημόσιο/Ιδιωτικό/Εσπεριν	Public/Private Secondary School (3 years), Secondary School (Evening Classes)	9		
	ISCED 3C shorter than 2 years	**4**					
	ISCED 3C 2 years or more	**5**					
	ISCED 3A-B	**6**					
	ISCED 3 (without distinction A-B-C, 2y+)	**7**	Λύκειο/ Τεχνική και Επαγγ	**High School/Vocational Technical School (day and night attendance for early school leavers-second chance schools)**	12	VET NOT POSSIBLE. CATEGORY MIXED.	
	ISCED 4C	**8**					
	ISCED 4A-B	**9**					
	ISCED 4 (without distinction A-B-C)	**10**					

CYPRUS*							
Int. Question No	**International English Version**	**Int. Value Code**	**National version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 5B	**11**	Τριτοβάθμια Μη-Πανεπιστ	**Non-University Degree/Diploma/Certificate leading to labour market, jobs at specific professional bodies i.e. policy, nursing, tourism, or ISCED5A**	14		
	ISCED 5A, bachelor degree	**12**	Πτυχίο Πανεπιστημίου	**Undergraduate degree**	16		
	ISCED 5A, master degree	**13**	Μεταπτυχιακό σε επίπεδο	**Postgraduate degree, Master's Degree-taught and research based**	18		
	ISCED 6	**14**	Διδακτορικό	**Doctorate**	21		

* Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

CZECH REPUBLIC							
Int. Question No	International English Version	Int. Value Code		English translation of the national version	Total years of schooling when level is completed	Vocational/ Academic	School starting age
B_Q01a	Which of the following qualifications is the highest you have obtained?		Jakého stupně vzdělání dosáhnete po ukončení Vašeho současného studia?	And which of these qualifications is the highest you have obtained?			6
	No formal qualification or below ISCED 1	**1**	Žádné formální vzdělání n	No formal education			
	ISCED 1	**2**	Dokončený první stupeň z	First level of basic school ISCED 1	5		
	ISCED 2	**3**	Základní vzdělání	basic ISCED 2	9		
	ISCED 3C shorter than 2 years	**4**	Vyučení bez maturity krat	vocational without maturita shorter than 2 years ISCED 3C shorter than 2 years	11	V	
	ISCED 3C 2 years or more	**5**	Vyučení bez maturity delš	vocational without maturita longer than 2 years ISCED 3C longer than 2 years	12	V	
	ISCED 3A-B	**6**	Vyučení s maturitou	ISCED 3A vocational with maturita	13	V	
			Střední odborné s maturit	ISCED 3A technical with maturita	13	V	
			Střední všeobecné s matu	ISCED 3A general with maturita	13	A	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**					
	ISCED 4C	**8**					
	ISCED 4A-B	**9**					
	ISCED 4 (without distinction A-B-C)	**10**	Středoškolská nástavba	ISCED 4 follow-up course	14	V	

CZECH REPUBLIC							
Int. Question No	**International English Version**	**Int. Value Code**		**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 5B	**11**	Vyšší odborné	ISCED 5B higher professional	16		
	ISCED 5A, bachelor degree	**12**	Bakalářské vysokoškolské	ISCED 5A, bachelor	16		
	ISCED 5A, master degree	**13**	Magisterské vysokoškolsk	ISCED 5A, master	18		
	ISCED 6	**14**	Postgraduální vzdělání	ISCED 6, post graduate	21		

DENMARK							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the qualifications on this card is the highest you have obtained?			Which of the following qualifications is the highest you have obtained?			6
	No formal qualification or below ISCED 1	**1**	Ingen offentligt godkendt eksamen eller under folkeskoleniveau	No formal education or below primary education			
	ISCED 1	**2**	Grundskole 1.-6. klasse	Primary school, grade 1-6	6		
	ISCED 2	**3**	Grundskole 7.-9(10). klasse	Lower secondary, grade 7-9(10)	9		
	ISCED 3C shorter than 2 years	**4**	Erhvervsfaglig uddannelse, under 2 år	Upper secondary vocational, less than 2 years	10	V	
	ISCED 3C 2 years or more	**5**	Erhvervsfaglig uddannelse, 2 år og derover	Upper secondary vocational, 2 years or more	12	V	
	ISCED 3A-B	**6**	Studentereksamen, HF, HHX, HTX	Upper secondary general, access to tertiary education	12	A	

DENMARK							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 3 (without distinction A-B-C, 2y+)	**7**	Erhvervsfaglig eller gymnasial uddannelse, 2 år og derover	Upper secondary undefined, 2 years or more	12	VET NOT POSSIBLE. CATEGORY MIXED.	
	ISCED 4C	**8**	Kort videregående erhvervsrettet uddannelse, under 2 år	Post secondary short programme, less than 2 years, lead to labour market	13	VET NOT POSSIBLE. CATEGORY MIXED.	
	ISCED 4A-B	**9**	Kort videregående studierettet uddannelse, under 2 år	Post secondary entrance course, access to tertiary education	13	A	
	ISCED 4 (without distinction A-B-C)	**10**	Kort videregående studierettet uddannelse, under 2 år	Post secondary non tertiary education, less than 2 years	13	VET NOT POSSIBLE. CATEGORY MIXED.	

DENMARK							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 5B	**11**	Kort/ mellemlang videregående uddannelse, 2 år og derover/ Professions bachelor, ikke forskningsbaseret	Tertiary not research based education, lead to labour market	15		
	ISCED 5A, bachelor degree	**12**	Bachelor grad	Bachelor degree	15		
	ISCED 5A, master degree	**13**	Kandidat eller master grad	Master degree	17		
	ISCED 6	**14**	Ph.d eller anden forskeruddannelse	Ph.d or otther research programme	20		
	Foreign qualification	**15**		Foreign qualification			

ENGLAND/NORTHERN IRELAND							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the qualifications on this card is the highest you have obtained?		And which of these was the highest qualification you have obtained?	Which of the following qualifications is the highest you have reached ?			5
	No formal qualification or below ISCED 1	**1**					
	ISCED 1	**2**					
	ISCED 2	**3**	No formal qualifications	No formal qualifications	11		
			Any other professional/ vocational qualifications/ apprenticeship	Any other professional/ vocational qualifications/ apprenticeship	11		
			Entry Level Qualifications	Entry Level Qualifications	11		
			Key Skills/ Basic Skills/ Essential Skills	Key Skills/ Basic Skills/ Essential Skills	11		
			YT Certificate/ YTP	YT Certificate/ YTP	11		
	ISCED 3C shorter than 2 years	**4**	City and Guilds (Level 1)	City and Guilds (Level 1)	11	V	
			RSA/ OCR (Level 1)	RSA/ OCR (Level 1)	11	V	
			National Qualifications (including SGA) (Scotland)	National Qualifications (including SGA) (Scotland)	11	A	
			Standard Grade or O Grade (Scotland)	Standard Grade or O Grade (Scotland)	11	A	
			Intermediate 1 or 2 NQs (Scotland)	Intermediate 1 or 2 NQs (Scotland)	11	A	
			O Level/GCSE/Vocational GCSE/CSE or equivalent	O Level/GCSE/Vocational GCSE/CSE or equivalent	11	A	
			GNVQ/ GSVQ (Level 1)	GNVQ/ GSVQ (Level 1)	11	V	
			NVQ/ SVQ (Level 1)	NVQ/ SVQ (Level 1)	11	V	

ENGLAND/NORTHERN IRELAND							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
			SCOTVEC, SCOTEC or SCOTBEC (Scotland)	SCOTVEC, SCOTEC or SCOTBEC (Scotland)	11	V	
			BTEC, BEC, TEC or EdExcel (Level 1)	BTEC, BEC, TEC or EdExcel (Level 1)	11	V	
	ISCED 3C 2 years or more	**5**	City and Guilds (Level 2)	City and Guilds (Level 2)	11	V	
			RSA/ OCR (Level 2)	RSA/ OCR (Level 2)	11	V	
			National Qualifications (including SGA) (Scotland)	National Qualifications (including SGA) (Scotland)	11	A	
			Standard Grade or O Grade (Scotland)	Standard Grade or O Grade (Scotland)	11	A	
			Intermediate 1 or 2 NQs (Scotland)	Intermediate 1 or 2 NQs (Scotland)	11	A	
			O Level/GCSE/Vocational GCSE/CSE or equivalent	O Level/GCSE/Vocational GCSE/CSE or equivalent	11	A	
			GNVQ/ GSVQ (Level 2)	GNVQ/ GSVQ (Level 2)	11	V	
			NVQ/ SVQ (Level 2)	NVQ/ SVQ (Level 2)	11	V	
			SCOTVEC, SCOTEC or SCOTBEC (Scotland)	SCOTVEC, SCOTEC or SCOTBEC (Scotland)	11	V	
			BTEC, BEC, TEC or EdExcel (Level 2)	BTEC, BEC, TEC or EdExcel (Level 2)	11	V	
	ISCED 3A-B	**6**	City and Guilds (Level 3)	City and Guilds (Level 3)	13	V	
			RSA/ OCR (Level 3)	RSA/ OCR (Level 3)	13	V	
			Advanced Highers or Certificate of 6th Year Studies (CSYS) (Scotland)	Advanced Highers or Certificate of 6th Year Studies (CSYS) (Scotland)	12	A	

ENGLAND/NORTHERN IRELAND							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
			AS level/ Vocational AS level or equivalent	AS level/ Vocational AS level or equivalent	12	A	
			GNVQ/ GSVQ (Level 3)	GNVQ/ GSVQ (Level 3)	13	V	
			NVQ/ SVQ (Level 3)	NVQ/ SVQ (Level 3)	13	V	
			Highers (Scotland)	Highers (Scotland)	12	A	
			A Level/ Vocational A Level or equivalent	A Level/ Vocational A Level or equivalent	13	A	
			SCOTVEC, SCOTEC or SCOTBEC (Scotland)	SCOTVEC, SCOTEC or SCOTBEC (Scotland)	13	V	
			BTEC, BEC, TEC or EdExcel (Level 3)	BTEC, BEC, TEC or EdExcel (Level 3)	13	V	
			ONC/OND (Level 3)	ONC/OND (Level 3)	13	V	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**					
	ISCED 4C	**8**					
	ISCED 4A-B	**9**					
	ISCED 4 (without distinction A-B-C)	**10**	Access to HE	Access to HE	13	A	
	ISCED 5B	**11**	RSA/ OCR (Level 4/5)	RSA/ OCR (Level 4/5)	15		
			NVQ/ SVQ (Level 4/5)	NVQ/ SVQ (Level 4/5)	15		
			Other Higher Education qualification below degree level	Other Higher Education qualification below degree level	15		
			SCOTVEC, SCOTEC or SCOTBEC (Scotland)	SCOTVEC, SCOTEC or SCOTBEC (Scotland)	15		
			BTEC, BEC, TEC or EdExcel (Level 4/5)	BTEC, BEC, TEC or EdExcel (Level 4/5)	15		
			HNC/HND (Level 4/5)	HNC/HND (Level 4/5)	15		
			Diploma in higher education	Diploma in higher education	15		
	ISCED 5A, bachelor degree	**12**					

ENGLAND/NORTHERN IRELAND							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 5A, master degree	**13**					
	ISCED 6	**14**					
	Foreign qualification	**15**	Foreign qualifications				
	Higher education (ISCED5A or ISCED6)	16	Nursing or other medical qualification not yet mentioned	Nursing or other medical qualification not yet mentioned	15		
			Teaching qualification (excluding PGCE)	Teaching qualification (excluding PGCE)	16		
			Degree level qualification including foundation degrees, graduate membership of a professional institute or PGCE, or higher	Degree level qualification including foundation degrees, graduate membership of a professional institute or PGCE, or higher	19		

					ESTONIA							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed (on average)**	**School starting age**		**Total years of schooling when level is completed (born before 1980, mother tonque Russian)**	**School starting age**		**Total years of schooling when level is completed (born 1980-1986, mother tonque Russian)**	**School starting age**
B_Q01a	Which of the following qualifications is the highest you have obtained?		Milline sellel kaardil nimetatuist on Teie kõrgeim haridustase?	Which of the following qualifications is the highest you have obtained?		7			7			6
	No formal qualification or below ISCED 1	**1**	Alghariduseta (vähem kui 3 klassi)	Without primary education								
	ISCED 1	**2**	Algharidus (3-6 klassi)	Primary education	6			3			4	
	ISCED 2	**3**	Põhiharidus (7-9 klassi)	Basic education	8			7			8	
			Kutseharidus (sisseastumisel ei nõutud põhiharidust)	(basic education not required at admission)	9.5			8.5			9.5	
	ISCED 3C shorter than 2 years	**4**	Kutseharidus põhihariduse baasil (õppekava nominaalkestus alla 2 aasta)	Vocational education on the basis of basic education (nominal time of studies less than 2 years)	9.5			8.5			9.5	

					ESTONIA							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed (on average)**	**School starting age**		**Total years of schooling when level is completed (born before 1980, mother tonque Russian)**	**School starting age**		**Total years of schooling when level is completed (born 1980-1986, mother tonque Russian)**	**School starting age**
	ISCED 3C 2 years or more	**5**	Kutseharidus põhihariduse baasil (õppekava nominaalkestus 2 aastat või enam)	Vocational education on the basis of basic education (nominal time of studies 2 years or more)	10.5			9.5			10.5	
	ISCED 3A-B	**6**	Üldkeskharidus	General secondary education	11			10			11	
			Kutsekeskharidus põhihariduse baasil	Vocational secondary education on the basis of basic education	11			10			11	
			Keskeriharidus põhihariduse baasil	Secondary specialised education on the basis of basic education	11			10			11	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**										
	ISCED 4C	**8**										

					ESTONIA							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed (on average)**	**School starting age**		**Total years of schooling when level is completed (born before 1980, mother tonque Russian)**	**School starting age**		**Total years of schooling when level is completed (born 1980-1986, mother tonque Russian)**	**School starting age**
	ISCED 4A-B	**9**	Kutsekeskharidus keskhariduse baasil	Vocational secondary education on the basis of secondary education	13			12			13	
	ISCED 4 (without distinction A-B-C)	**10**										
	ISCED 5B	**11**	Keskeriharidus keskhariduse baasil	Secondary specialised education on the basis of secondary education	14			13			14	
			Kutsekõrgharidus, diplomiõpe, rakenduskõrgharidus	Applied higher education	15.5			13.5			14.5	
	ISCED 5A, bachelor degree	**12**	Bakalaureusekraad (3+2 süsteemi järgi, alustatud pärast 2002.a)	Bachelor's degree (3+2 system, started after 2002)	15			13			14	
	ISCED 5A, master degree	**13**	Bakalaureusekraad (4+2 süsteemi järgi, alustatud enne 2002.a)	Bachelor's degree (4+2 system, started before 2002)	15			14			15	

					ESTONIA							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed (on average)**	**School starting age**		**Total years of schooling when level is completed (born before 1980, mother tonque Russian)**	**School starting age**		**Total years of schooling when level is completed (born 1980-1986, mother tonque Russian)**	**School starting age**
			Enne 1992. aastat alustatud kõrgharidus (diplomeeritud spetsialistiõpe)	Higher education (studies for diploma of specialist), started before 1992	16			15			16	
			Magistrikraad (3+2 süsteemi järgi, sh integreeritud bakalaureuse- ja magistriõpe)	Master's degree (3+2 system, incl integrated Bachelor and Master's studies)	17			15			16	
			Magistrikraad (4+2 süsteemi järgi)	Master's degree (4+2 system)	17			16			17	
	ISCED 6	**14**	Doktorikraad (sh kandidaadikraad)	Doctoral degree (incl Candidate of Doctor)	21			20			21	
	Foreign qualification	**15**	Välisriigis omandatud haridus, palun täpsusta	Foreign qualification, please specify								

						ESTONIA					
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**		**Total years of schooling when level is completed (born before 1983, mother tonque Estonian)**	**School starting age**		**Total years of schooling when level is completed (Estonian mother tonque, born 1983 or after; Russian mother tonque, born 1987 or after)**	**School starting age**	**Vocational/ General**
B_Q01a	Which of the following qualifications is the highest you have obtained?		Milline sellel kaardil nimetatuist on Teie kõrgeim haridustase?	Which of the following qualifications is the highest you have obtained?			7			6	
	No formal qualification or below ISCED 1	**1**	Alghariduseta (vähem kui 3 klassi)	Without primary education							
	ISCED 1	**2**	Algharidus (3-6 klassi)	Primary education		3			6		
	ISCED 2	**3**	Põhiharidus (7-9 klassi)	Basic education		8			9		
			Kutseharidus (sisseastumisel ei nõutud põhiharidust)	(basic education not required at admission)		9.5			9.5		
	ISCED 3C shorter than 2 years	**4**	Kutseharidus põhihariduse baasil (õppekava nominaalkestus alla 2 aasta)	Vocational education on the basis of basic education (nominal time of studies less than 2 years)		9.5			10.5		V

						ESTONIA					
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**		**Total years of schooling when level is completed (born before 1983, mother tonque Estonian)**	**School starting age**		**Total years of schooling when level is completed (Estonian mother tonque, born 1983 or after; Russian mother tonque, born 1987 or after)**	**School starting age**	**Vocational/ General**
	ISCED 3C 2 years or more	**5**	Kutseharidus põhihariduse baasil (õppekava nominaalkestus 2 aastat või enam)	Vocational education on the basis of basic education (nominal time of studies 2 years or more)		10.5			11.5		V
	ISCED 3A-B	**6**	Üldkeskharidus	General secondary education		11			12		A
			Kutsekeskharidus põhihariduse baasil	Vocational secondary education on the basis of basic education		11			12		V
			Keskeriharidus põhihariduse baasil	Secondary specialised education on the basis of basic education		11			12		V
	ISCED 3 (without distinction A-B-C, 2y+)	**7**									
	ISCED 4C	**8**									

						ESTONIA					
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**		**Total years of schooling when level is completed (born before 1983, mother tonque Estonian)**	**School starting age**		**Total years of schooling when level is completed (Estonian mother tonque, born 1983 or after; Russian mother tonque, born 1987 or after)**	**School starting age**	**Vocational/ General**
	ISCED 4A-B	**9**	Kutsekeskharidus keskhariduse baasil	Vocational secondary education on the basis of secondary education		13			14		V
	ISCED 4 (without distinction A-B-C)	**10**									
	ISCED 5B	**11**	Keskeriharidus keskhariduse baasil	Secondary specialised education on the basis of secondary education		14			15		
			Kutsekõrgharidus, diplomiõpe, rakenduskõrgharidus	Applied higher education		14.5			15.5		
	ISCED 5A, bachelor degree	**12**	Bakalaureusekraad (3+2 süsteemi järgi, alustatud pärast 2002.a)	Bachelor's degree (3+2 system, started after 2002)		14			15		
	ISCED 5A, master degree	**13**	Bakalaureusekraad (4+2 süsteemi järgi, alustatud enne 2002.a)	Bachelor's degree (4+2 system, started before 2002)		15			16		

						ESTONIA					
Int. Question No	International English Version	Int. Value Code	National version (text)	English translation of the national version		Total years of schooling when level is completed (born before 1983, mother tonque Estonian)	School starting age		Total years of schooling when level is completed (Estonian mother tonque, born 1983 or after; Russian mother tonque, born 1987 or after)	School starting age	Vocational/ General
			Enne 1992. aastat alustatud kõrgharidus (diplomeeritud spetsialistiõpe)	Higher education (studies for diploma of specialist), started before 1992		16					
			Magistrikraad (3+2 süsteemi järgi, sh integreeritud bakalaureuse- ja magistriõpe)	Master's degree (3+2 system, incl integrated Bachelor and Master's studies)		16			17		
			Magistrikraad (4+2 süsteemi järgi)	Master's degree (4+2 system)		17			18		
	ISCED 6	**14**	Doktorikraad (sh kandidaadikraad)	Doctoral degree (incl Candidate of Doctor)		21			21		
	Foreign qualification	**15**	Välisriigis omandatud haridus, palun täpsusta	Foreign qualification, please specify							

FINLAND							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the qualifications on this card is the highest you have obtained?		Mikä on korkein suorittamanne tutkinto?	Which of the following qualifications is the highest you have obtained?			7
	No formal qualification or below ISCED 1	**1**	Ei muodollista tutkintoon johtavaa koulutusta	No formal education			
	ISCED 1	**2**	Peruskoulun luokat 1-6, kansakoulu, osa keskikoulua	Grades 1-6 of comprehensive school, primary school, part of middle school (ISCED 1)	6		
	ISCED 2	**3**	Peruskoulun luokat 7-9(10), keskikoulu	Grades 7-9(10) of comprehensive school, middle school (ISCED 2)	9		
	ISCED 3C shorter than 2 years	**4**					
	ISCED 3C 2 years or more	**5**					
	ISCED 3A-B	**6**					
	ISCED 3 (without distinction A-B-C, 2y+)	**7**	Ammatillinen kouluasteen tutkinto, ammatillinen perustutkinto, ammattitutkinto	Upper secondary vocational education and training (ISCED 3)	11	V	
			Lukio	General upper secondary school (ISCED 3)	12	A	
	ISCED 4C	**8**					
	ISCED 4A-B	**9**					
	ISCED 4 (without distinction A-B-C)	**10**	Erikoisammattitutkinto	Specialist vocational qualification (ISCED 4)	12	V	

FINLAND							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 5B	**11**	Ammatillinen opistoasteen tutkinto	Vocational post-secondary qualification (ISCED 5B)	14		
	ISCED 5A, bachelor degree	**12**	Ammattikorkeakoulututkinto	Polytechnic degree (ISCED 5A)	16		
			Alempi korkeakoulututkinto, kandidaatin tutkinto	Bachelor's degree (ISCED 5A)	15		
	ISCED 5A, master degree	**13**	Ylempi korkeakoulututkinto, maisterin tutkinto, ylempi ammattikorkeakoulututkinto	Master's degree (ISCED 5A)	17		
	ISCED 6	**14**	Lisensiaatin ja tohtorin tutkinnot	Licentiate's and doctor's degrees (ISCED 6)	Licentiate 19 Doctor 21		

FLANDERS (BELGIUM)									
Int. Question No	**International English Version**	**Int. Value Code**	**National version**	**English translation of the national version**	**Recoding suggestion by DPC**	**Recoding instruction by country**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the following qualifications is the highest you have obtained?			What is the highest level of education you have ever successfully completed?	**B_Q01aBE --> B_Q01a 1 --> 1 2 --> 2 3 --> 3 4 --> 5 5 --> 6 6 --> 7 7 --> 9 8 --> 11 9 --> 12 10 --> 13 11 --> 14 12 --> 15**				6
	No formal qualification or below ISCED 1	**1**	Geen onderwijs of het lager onderwijs niet beëindigd	No education or not completed primary education		Use international Response category 1			
	ISCED 1	**2**	Lager onderwijs of basiseducatie	Primary education or adult basic education		Use international Response category 2	6		

FLANDERS (BELGIUM)									
Int. Question No	**International English Version**	**Int. Value Code**	**National version**	**English translation of the national version**	**Recoding suggestion by DPC**	**Recoding instruction by country**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 2	**3**	Lager secundair onderwijs (of eerste graad)	Lower secondary education (or first stage secondary education)		Use international Response category 3	8		
	ISCED 3C shorter than 2 years	**4**		No correspondence					
	ISCED 3C 2 years or more	**5**	Volledig beroepssecundair onderwijs	Vocational secondary education		Use international Response category 5	12	VET NOT POSSIBLE, CATEGORY MIXED	
	ISCED 3A-B	**6**	Volledig algemeen, technisch of kunst secundair onderwijs	General or technical secondary education		Use international Response category 6	12		
	ISCED 3 (without distinction A-B-C, 2y+)	**7**	Hoger secundair onderwijs (geen onderwijsvorm)	Upper secondary education (no education form)		Use international Response category 7	12		
	ISCED 4C	**8**		No correspondence					

FLANDERS (BELGIUM)									
Int. Question No	**International English Version**	**Int. Value Code**	**National version**	**English translation of the national version**	**Recoding suggestion by DPC**	**Recoding instructio n by country**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 4A-B	**9**	Voortgezet secundair onderwijs dat toegang geeft tot hoger onderwijs (vierde graad of derde jaar van de derde graad van het secundair onderwijs)	Post-secondary non-tertiary education giving access to higher education (4th stage or 3rd year of 3rd stage secondary education)		Use internation al Response category 9	13		
	ISCED 4 (without distinction A-B-C)	**10**		No correspondence			13		
	ISCED 5B	**11**	Hoger onderwijs van één cyclus (korte type / professionele bacheloropleiding)	1 cycle higher education (short type / professional bachelor courses)		Use internation al Response category 11	15		
	ISCED 5A, bachelor degree	**12**	Academische bacheloropleiding (universitaire kandidatuuropleiding)	Academic bachelor courses (University candidate degree)		Use internation al Response category 12	15		
	ISCED 5A, master degree	**13**	Universitair onderwijs of hoger onderwijs van twee cycli (lange type / masteropleiding)	University education or 2 cycle higher education (long type / master courses)		Use internation al Response category 13	16		

FLANDERS (BELGIUM)									
Int. Question No	**International English Version**	**Int. Value Code**	**National version**	**English translation of the national version**	**Recoding suggestion by DPC**	**Recoding instructio n by country**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 6	**14**	doctoraat	Doctorate		Use internation al Response category 14	20		
						Use internation al Response category 15			

FRANCE							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the qualifications on this card is the highest you have obtained?						6
	No formal qualification or below ISCED 1	**1**	Aucun ou enseignement primaire inachevé	Never been to school or incomplete primary	5		
	ISCED 1	**2**	École primaire	Primary school	5		
	ISCED 2	**3**	Aucun diplôme (No diploma)	No diploma	5		
			Certificat d'études primaires (CEP) ou équivalent	Primary school certificate	5		
			Brevet des collèges ou équivalent	Secondary education, 1st cycle diploma like "brevet des collèges"	9		
	ISCED 3C shorter than 2 years						
	ISCED 3C 2 years or more	**4**	CAP, BEP ou diplôme de ce niveau	Vocational training diploma like "CAP" or "BEP"	11	V	
		5	Brevet professionnel (BP, BPA) ou de technicien (BT, BTA) ou diplôme de ce niveau	Professional or technical "brevet" or similar diploma	14	V	
	ISCED 3A-B	**6**	Baccalauréat professionnel	Professional baccalauréat	13	V	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**	Baccalauréat technologique	Technological baccalauréat	12	A	
			Baccalauréat général (General baccalauréat)	General baccalauréat	12	A	
	ISCED 4C	**8**					
	ISCED 4A-B	**9**					
	ISCED 4 (without distinction A-B-C)	**10**					
	ISCED 5B	**11**	BTS, DUT, DEUST, Diplôme des professions sociales et de la santé de niveau bac+2 (infirmière, kinésithérapeute, assistante sociale…)	Vocational training and technical diplomas up to Bac+2, BTS, DUT, DEUST, diplomas related to healthcare up to Bac+2	14		

FRANCE							
Int. Question No	International English Version	Int. Value Code	National Version	English translation of the national version	Total years of schooling when level is completed	Vocational/ Academic	School starting age
	ISCED 5A, bachelor degree	**12**	Diplôme de 1er cycle universitaire, DEUG, DUEL, DUES, PCEM, certificat d'aptitude pédagogique, certificat de fin d'études normales (CFEN)	University education, 1st cycle - DEUG, DUEL, DUES, PCEM or other diplomas	14		
			Diplôme de 2ème cycle universitaire : Licence, maîtrise, IUFM, CAPE, CAPES, CAPET, agrégation…	University education, 2nd cycle diplomas like licence, maîtrise, IUFM, CAPE, CAPES, CAPET, agrégation…	15		
	ISCED 5A, master degree	**13**	Diplôme d'une grande école (ingénieur, commerce…), diplôme d'études comptables supérieures (DECS), d'avocat (CAPA), d'expert-comptable, ingénieur du CNAM,…	Higher engineering school, higher business school, expert accounting qualification, lawyer qualification	17		
			University education, 3rd cycle (DES, DEA, DESS (=MPHIL), master) ; thesis and doctorate (=PHD) related to healthcare	University education, 3rd cycle (DES, DEA, DESS (=MPHIL), master) ; thesis and doctorate (=PHD) related to healthcare	17		
	ISCED 6	**14**	Thesis and doctorate (=PHD) NOT related to healthcare	Thesis and doctorate (=PHD) NOT related to healthcare	20		

GERMANY							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the qualifications on this card is the highest you have obtained?		Welchen höchsten allgemein bildenden Schulabschluss haben Sie? Bitte sagen Sie es mir anhand dieser Liste.	What is the highest general education school leaving certificate that you hold? Please tell me according to this list.			6
	No formal qualification or below ISCED 1	**1**	Von der Schule abgegangen ohne Hauptschulabschluss (Volksschulabschluss), aber nach Beendigung der Grundschule	No formal education or left school without completing primary school grades			
	ISCED 1	**2**	Von der Schule abgegangen ohne Hauptschulabschluss (Volksschulabschluss)	Completed primary school grades, but left school without a Hauptschulabschluss (general education school leaving certificate obtained on completion of grade 9) or a leaving certificate from the Volksschule (the former name for compulsory school)	7		

GERMANY							
Int. Question No	International English Version	Int. Value Code	National version (text)	English translation of the national version	Total years of schooling when level is completed	Vocational/ Academic	School starting age
	ISCED 2	**3**	Hauptschulabschluss (Volksschulabschluss)	Hauptschulabschluss (general education school leaving certificate obtained on completion of grade 9 at a Hauptschule or any other lower secondary level school) or a leaving certificate from the Volksschule (the former name for compulsory school)	9		
			Realschulabschluss (Mittlere Reife)	Realschulabschluss (general education school leaving certificate obtained on completion of grade 10 at a Realschule or, under certain circumstances, at other lower secondary level school types. It can also be obtained at a later stage during vocational training at upper secondary level).	10		

GERMANY							
Int. Question No	International English Version	Int. Value Code	National version (text)	English translation of the national version	Total years of schooling when level is completed	Vocational/ Academic	School starting age
			Abgang von der Polytechnischen Oberschule nach der 8. Klasse nach 1965	Left the Polytechnische Oberschule (Polytechnical High School, main secondary school type in former GDR) after 8th grade after 1965.	8		
			Abgang von der Polytechnischen Oberschule nach der 10. Klasse (vor 1965: 8. Klasse)	Left the Polytechnische Oberschule (Polytechnical High School, main secondary school type in former GDR) after10th grade (pre 1965: 8th grade)	10		
	ISCED 3C shorter than 2 years	**4**	n/a				
	ISCED 3C 2 years or more	**5**	n/a				
	ISCED 3A-B	**6**	Fachhochschulreife, Abschluss Fachoberschule	Fachhochschulreife, a qualification obtained at a Fachoberschule (vocational school at upper secondary level) after 12 years of schooling. It entitles the holder to study at a Fachhochschule (technical college/university of applied sciences).	12	A	

GERMANY							
Int. Question No	International English Version	Int. Value Code	National version (text)	English translation of the national version	Total years of schooling when level is completed	Vocational/ Academic	School starting age
			Allgemeine oder fachgebundene Hochschulreife/ Abitur (Gymnasium bzw. EOS, EOS mit Lehre)	General higher education entrance qualification entitling holder to study all subjects at a higher education institution, or a discipline-specific qualification entitling the holder to study only certain subjects	13	A	
			Beruflich-betriebliche Berufsausbildung (Lehre)	(Completed) apprenticeship in the dual system (combination of in-company training and training at vocational school at upper secondary level)	13	V	
			Beruflich-schulische Ausbildung (Berufsfachschule, Handelsschule, Kollegschule oder Schule des Gesundheitswesens (1-jährig)	Basic vocational training at a Berufsfachschule (full-time vocational school at upper secondary level), Handelsschule (commercial college), Kollegschule (vocational college) or a school for medical assistants (1-year course)	12	V	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**	n/a				
	ISCED 4C	**8**	n/a				

GERMANY							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 4A-B	**9**	Abitur oder Hochschulreife an der Abendschule	General higher education entrance qualification at evening school	14	V	
	ISCED 4 (without distinction A-B-C)	**10**	(Fach)Hochschulreife + berufliche Ausbildung	Higher education entry qualification but did not go to higher education but completed apprenticeship	14	V	
	ISCED 5B	**11**	**Meister, Techniker oder gleichwertiger Fachschulabschluss**	**Trade and Technical school**	15		
			Berufs- oder Fachakademie, Schule des Gesundheitswesens (2- bis 3-jährig)	*Specialised academy, Vocational Academy, Health Sector School (2 - 3 years)*	15		
	ISCED 5A, bachelor degree	**12**	Fachhochschulabschluss (Bachelor)	Bachelor's degree from a Fachhochschule (university of applied sciences/technical college)	16		
			Hochschulabschluss (Bachelor)	(here) a Bachelor's degree from a university (as opposed to a university of applied sciences)	16		
	ISCED 5A, master degree	**13**	Fachhochschulabschluss (Master; Diplom)	Master's or Diplom degree from a Fachhochschule (university of applied sciences/technical college)	17		

GERMANY							
Int. Question No	International English Version	Int. Value Code	National version (text)	English translation of the national version	Total years of schooling when level is completed	Vocational/ Academic	School starting age
			Hochschulabschluss (Diplom. Magister, Staatsexamen; Master)	(Here) an advanced degree (Diplom, Magister, Master's, or State Examination in medicine, teaching or law) from a university as opposed to a university of applied sciences	18		
	ISCED 6	**14**	Promotion	doctorate	21		

IRELAND							
Int. Question No	International English Version	Int. Value Code	National Version	English translation of the national version	Total years of schooling when level is completed	Vocational/ Academic	School starting age
B_Q01a	Which of the qualifications on this card			What is the highest level of edcuation or training			4
	No formal qualification or below ISCED 1	**1**		No formal education or training \ Pre-primary education (or new FETAC certificate at NFQ level 1)			
	ISCED 1	**2**		Primary education (or FETAC Certificate at NFQ level 2)	8		
	ISCED 2	**3**		Secondary 1 (Junior/Inter/Group Certificate, NCVA foundation Certificate, FAS IAS Certificate or FETAC Certificate at NFQ level 3)	11		
				Transition year programme	12		
	ISCED 3C shorter than 2 years	**4**					
	ISCED 3C 2 years or more	**5**					
	ISCED 3A-B	**6**					
	ISCED 3 (without distinction A-B-C, 2y+)	**7**		Secondary 2 (Leaving Certificate, traditional, vocational applied)	14	A	

IRELAND							
Int. Question No	International English Version	Int. Value Code	National Version	English translation of the national version	Total years of schooling when level is completed	Vocational/ Academic	School starting age
	ISCED 4C	**8**		Technical or Vocational (e.g. Secretarial courses, Certificate in Hotel Operations, PLCs, FAS National Skills/Specific Skills Certificate or FETAC Certificate at NFQ level 4 & 5)	15	V	
				Advanced Certificate (Completed apprenticeships, Teagasc Farming or Horticulture Certificate, National Craft Certificate or FETAC Advanced certificate at NFQ level 6)	18	V	
	ISCED 4A-B	**9**					
	ISCED 4 (without distinction A-B-C)	**10**					
	ISCED 5B	**11**		Higher Certificate (e.g. National Certificate (NCEA/DIT/IOT), Cadetship (army, air corps or naval service), Diploma in Police Studies or HETAC/DIT Higher Certificate at NFQ level 6)	16		

IRELAND							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
				Diploma (e.g. National Diploma (HETAC/NCEA), Bachelor Degree (DIT), 3 year Diploma or new Ordinary Bachelor Degree at NFQ level 7)	17		
	ISCED 5A, bachelor degree	**12**		Honours Bachelor Degree, Graduate Diploma *(or Higher Diploma at NFQ level 8)*	18		
				Professional (Honours Bachelor Degree equivalent or higher)	18		
	ISCED 5A, master degree	**13**		Post-Graduate (e.g. Post Graduate Diploma at NFQ level 9, Masters Degree (taught or researched) at NFQ level 9)	19		
	ISCED 6	**14**		Doctorate or higher (e.g. Doctoral Degree/higher Doctorate at NFQ level 10)	21		
	Foreign qualification	**15**		Foreign qualification	N\A		
B_Q02b							

ITALY							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the following qualifications is the highest you have obtained?		Quale dei seguenti è il titolo di studio più alto che ha ottenuto?	Which of the following qualifications is the highest you have obtained?			6
	No formal qualification or below ISCED 1	**1**	Nessun titolo o meno della licenza elementare (ISCED 0)	Non formal education or below ISCED 1	0		
	ISCED 1	**2**	Licenza elementare (ISCED 1)	Primary education or first stage of basic education	5		
	ISCED 2	**3**	Licenza media e nuovo obbligo (ISCED 2)	Lower secondary or second stage of basic education	8		
	ISCED 3C shorter than 2 years	**4**	Corsi regionali brevi (I livello) - (ISCED 3C shorter than 2 years)	Regional Vocational training qualification 1st level	9		
	ISCED 3C 2 years or more	**5**	Qualifica degli istituti professionali di Stato (ISCED 3C 2 years or more)	Educational and vocational training qualification	11		
	ISCED 3A-B	**6**	Diploma quinquennale (ISCED 3A)	Upper secondary education	13	VET NOT POSSIBLE. CATEGORY MIXED	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**					
	ISCED 4C	**8**					
	ISCED 4A-B	**9**					

ITALY							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 4 (without distinction A-B-C)	**10**	Corso post-diploma non accademico (IFTS e corsi regionali di II livello) – (ISCED 4)	Post-secondary non tertiary education (Regional vocational training qualification 2nd level or Certification of higher technical specialisation)	15	V	
	ISCED 5B	**11**	Diploma di Conservatorio di musica, di Accademia di belle arti, di Accademia di danza, di Attore o Regista o ISIA (ISCED 5B)	Music Conservatory Diploma or National Dance Academy Diploma or Diploma of actor or director	16	V	
	ISCED 5A, bachelor degree	**12**	Laurea di 3-5-6 anni (compreso Vecchio Ordinamento) o Diploma accademico (ISCED 5A)	First stage of tertiary education (Universitary Diploma or Laurea degree or Second level degree).	18	N.A.	
	ISCED 5A, master degree	**13**	Corsi post-laurea (ISCED 5A) OR Corso di specializzazione post-laurea (di almeno 2 anni) – (ISCED 5A)	First or second level post-lauream master degree or specialisation degree (ISCED 5A)	19	N.A.	
	ISCED 6	**14**	Dottorato di ricerca (ISCED 6)	Research Doctoral degree	21	N.A.	
	Foreign qualification	**15**	12. Titolo rilasciato all'estero (specificare)	Foreign qualification, please specify		-	

JAPAN								
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**	**Country comment**
B_Q01a	**Which of the qualifications on this card is the highest you have obtained?**		**あなたの最終学歴をカードの中から選んでください。**	Which of the following qualifications is the highest you have obtained?			6	JPN NPM would like to stress that she considers the JPN educational system a single track, rather than a general/vocational track system
	No formal qualification or below ISCED 1	**1**	学校には行ったことがない、 または小学校中退	No formal school education, Dropped out of elementary school				
	ISCED 1	**2**	小学校	Elementary school, Special education school (elementary department)	6			
	ISCED 2	**3**	中学校	Lower secondary school, Secondary education school (lower division), Special education school (lower secondary department)	9			
	ISCED 3C shorter than 2 years	**4**	高校の別科	Short-term course of upper secondary school, Short-term course of secondary education school (upper division), Short-term course of special education school (upper secondary department)	10	V		

JAPAN								
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**	**Country comment**
	ISCED 3C 2 years or more	**5**	高校の専門学科、 専修学校高等課程	(Full day / day / evening / corresponding) Specialized course of upper secondary school, (Full day / day / evening) Specialized course of secondary education school (upper division), Specialized course of special education school (upper secondary department), Specialized training college (upper secondary course)	12	V		

JAPAN								
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**	**Country comment**
	ISCED 3A-B	**6**	高校の普通科または総合学科 高等専門学校（第1-3学年）	General / integrated course of Upper Secondary school (including correspondence course), General / integrated course of secondary education school (upper division), General course of special education school (upper secondary department), College of technology (1st-3rd year)	12	A		
	ISCED 3 (without distinction A-B-C, 2y+)	**7**	高卒認定合格者（旧大検合格者を含む）	Passed Upper Secondary School Equivalency Examination, Unknown	9			
	ISCED 4C	**8**						
	ISCED 4A-B	**9**						

JAPAN								
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**	**Country comment**
	ISCED 4 (without distinction A-B-C)	**10**	高校の専攻科、 短期大学または大学の別科	Advanced course of upper secondary school, Advanced course of secondary education school (upper division), Advanced course of special education school (upper secondary department), Short-term course of junior college, Short-term course of univerity	13	V		
	ISCED 5B	**11**	短期大学、高等専門学校（第4-5学年）、 短期大学または高等専門学校の専攻科 専門学校（専修学校専門課程）	Regular course of junior college (including correspondence course), Advanced course of junior college, Regular course of college of technology, Advanced course of college of technology, Specialised training college（post-secondary course)	14			

JAPAN								
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**	**Country comment**
	ISCED 5A, bachelor degree	**12**	大学学士課程、 大学の専攻科	Undergraduate programs of University (including correspondence course), Advanced course of university	16			
	ISCED 5A, master degree	**13**	大学院修士課程または博士前期課程、 大学院専門職学位課程（法科大学院を含む）	Master's programs / Doctoral programs (lower division) of university (including correspondence course), Professional degree's programs of university (including correspondence course), Professional degree's programs of graduate law school	18			
			博士課程満期退学	Completed all work of doctoral program except doctoral thesis	21			
	ISCED 6	**14**	大学院博士課程	Doctoral programs of university (including correspondence course)	21			
			専修学校一般課程、 各種学校	Specialised training college (general course), Miscellaneous school	9			

KOREA							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the following qualifications is the highest you have obtained?		] 보기 카드 1을 보십시오. 귀하의 최종 학력은 무엇입니까?	What is the highest level of formal education you have ever successfully completed?			6
	No formal qualification or below ISCED 1	**1**	무학	no formal education or below Elementary			
	ISCED 1	**2**	초졸	Elementary	6		
	ISCED 2	**3**	중졸	Middle School	9		
	ISCED 3C shorter than 2 years	**4**					
	ISCED 3C 2 years or more	**5**	고졸(전문계/ 이전의 실업계)	High School(vocational education)	12	V	
	ISCED 3A-B	**6**	고졸(일반계)	High School(college prep.)	12	A	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**					
	ISCED 4C	**8**					
	ISCED 4A-B	**9**					
	ISCED 4 (without distinction A-B-C)	**10**					
	ISCED 5B	**11**		Master's degree(Specialized(vocational) graudate schools)			
			2-3년제 전문대 졸	2-3 year college	16		
				4 year college of education(Bachelor degree)			

KOREA							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 5A, bachelor degree	**12**	일반 4년제 대학교 졸(학사) 특수 4년제대학(교육대학, 산업대학, 경찰대학 등) 졸(학사)	4 year university(Bachelor degree)	16		
	ISCED 5A, master degree	**13**	일반대학원 석사 학위취득 특수대학원 석사 학위취득 전문대학원 석사 학위취득	Master's degree(general univeristies)	18		
	ISCED 6	**14**	박사 학위취득	Doctoral degree	22		

NETHERLANDS							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the following qualifications is the highest you have obtained?		De volgende vraag gaat over het hoogste onderwijsniveau dat u volledig heeft afgemaakt. Kunt u aangeven welk niveau dat was. We bedoelen hietr onderwijs dat tot een echt schooldiploma leidt, zoals mavo, mbo of universiteit.	The next question is about the highest level of (formal) education that you completed entirely. Can you tell what level that was?			
	INTERVIEWER: If the respondent is currently enrolled in an educational programme, emphasize that the question refers to education that has been completed, and that current education will be addressed in a later question.			*INTERVIEWER: If the respondent is currently enrolled in an educational programme, emphasize that the question refers to education that has been completed, and that current education will be addressed in a later question.*			
	No formal qualification or below ISCED 1	**1**	geen diploma	No formal qualification or below ISCED 1			
	ISCED 1	**2**	basisonderwijs, lagere school, speciaal lager onderwijs	primary education (isced 1, piaac 2)	7		

NETHERLANDS							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 2	**3**	vmbo praktijkonderwijs, ibo, ivbo, speciaal voortgezet onderwijs	secondary education, first cycle, middle (isced 2c, piaac 3)	11		
			vmbo (bl, kl), lbo, vbo, bol/mbo 1 jarig (assistentenopleiding), kmbo 1 jarig, bbl 1 jarig	secondary education, first cycle, middle (isced 2, piaac 3)	11		
			mulo, mavo, vmbo (tl, gl)	secondary education, first cycle, high (isced 2b, piaac 3)	11		
	ISCED 3C shorter than 2 years	**4**					
	ISCED 3C 2 years or more	**5**	leerlingwezen primair, bbl 2 jarig	secondary education, first cycle, middle (isced 2c, piaac 5)	13	V	
			bol/mbo 2 jarig, kmbo 2 jarig	secondary education, second cycle, low (isced 3c, piaac 5)	13	V	
			leerlingwezen secundair of tertiair, bbl 3- of 4-jarig	secondary education, second cycle, middle (isced 3c, piaac 5)	14	V	
	ISCED 3A-B	**6**	bol/mbo 3 of 4 jarig	secondary education, second cycle, high (isced 3a, piaac 6)	14	V	
			havo, mms	secondary education, second cycle, middle (isced 3a, piaac 6)	12	A	
			vwo, gymnasium, hbs	secondary education, second cycle, high (isced 3a, piaac 6)	13	A	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**					
	ISCED 4C	**8**					

NETHERLANDS							
Int. Question No	International English Version	Int. Value Code	National version (text)	English translation of the national version	Total years of schooling when level is completed	Vocational/ Academic	School starting age
	ISCED 4A-B	**9**					
	ISCED 4 (without distinction A-B-C)	**10**					
	ISCED 5B	**11**	kort hbo, associate degree	tertiary education, first cycle, low (isced 5b, piaac 11)	14		
	ISCED 5A, bachelor degree	**12**	hbo 4 jarig, hbo bachelor	tertiary education, first cycle, middle (isced 5a, piaac 12)	16		
			universiteit bachelor	tertiary education, first cycle, high (isced 5a, piaac 12)	16		
	ISCED 5A, master degree	**13**	universiteit doctoraal, hbo/universiteit master	tertiary education, second cycle (isced 5a, piaac 13)	17		
	ISCED 6	**14**	doctoraat, medisch specialist	tertiary education, third cycle (isced 6, piaac 14)	21		

NORWAY							
Int. Question No	**International English Version**	**Int. Value Code**		**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the following qualifications is the highest you have obtained?			What is the highest education you have obtained?			5
	No formal qualification or below ISCED 1	**1**	Ingen formell utdanning	No formal education			
	ISCED 1	**2**	Barneskole	Primary school	7		
	ISCED 2	**3**	Ungdomsskole, folkeskole	Compulsory school. General education school leaving certificate obtained on completion of grade 10 (or any other lower secondary level school) or a leaving certificate from Folkeskole (the former name for compulsory school)	10		
	ISCED 3C shorter than 2 years	**4**	Gymnas, videregående utdanning, realskole, folkehøyskole to år eller mindre, ikke direkte studiekompetansegivende	Vocational/high school /folk high school education 2 years or shorter, not giving direct access to ISCED 5 level	12	V	
	ISCED 3C 2 years or more	**5**	Videregående utdanning som gir fagbrev, svennebrev eller tilsvarende yrkesfaglig kompetanse. 2 års varighet eller mer.	Vocational education with craft certificate, no direct access to ISCED 5 level.	14	V	
	ISCED 3A-B	**6**	Gymnas, realskole eller videregående opplæring som gir generell studiekompetanse	Vocational/high school education giving direct access to ISCED 5 leve, 3 years or shorter.	13	A	

NORWAY							
Int. Question No	**International English Version**	**Int. Value Code**		**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 3 (without distinction A-B-C, 2y+)	**7**					
	ISCED 4C	**8**	Fagskoleutdanning og annen yrkesrettet påbygging til videregående opplæring	2 years education at high school or supplementary education for adults giving access to ISCED 5 level	15	V	
	ISCED 4A-B	**9**	Forkurs til universitet og høgskole som ikke gir vekttall/studiepoeng	Introductory course to provide direct access to college/university. pre-degree foundation courses or short vocational programmes	14	A	
	ISCED 4 (without distinction A-B-C)	**10**					
	ISCED 5B	**11**	Toårig høgskolekandidatgrad	Education at college/university or other post secondary education 2 years or shorter	15		
	ISCED 5A, bachelor degree	**12**	Bachelor, cand. Mag. eller annen universitets- og høgskoleutdanning, tilsvarende inntil fire års heltidsstudier (80 vekttall/240 studiepoeng eller mindre)	Education at college/university or other post secondary education 3 years or shorter	16		
	ISCED 5A, master degree	**13**	Master, hovedfag eller annen universitets- og høgskoleutdanning, tilsvarende mer enn fire års heltidsstudier (mer enn 80 vekttall/240 studiepoeng)	Education at college/university or other post secondary education, 4 years or longer	18		
	ISCED 6	**14**	Forskerutdanning	Second stage of tertiary education (post graduate)	21		

POLAND											
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed, born before 1952**	**School starting age**	**Total years of schooling when level is completed, born after 1951 and before 1986**	**School starting age**	**Total years of schooling when level is completed, born after 1985**	**School starting age**	**Vocational/ Academic**
B_Q01a	Which of the following qualifications is the highest you have obtained?		Proszę na tej karcie wskazać najwyższy poziom wykształcenia, jakie Pan posiada.	Which of the following qualifications is the highest you have obtained?		7		7		7	
	No formal qualification or below ISCED 1	**1**	Niepełne podstawowe	incomplete primary							
	ISCED 1	**2**	Podstawowe I (po reformie) (ISCED 1)	primary I ISCED 1	6		6		6		
	ISCED 2	**3**	podstawowe II (przed reformą lub gimnazjum) (ISCED 2)	primary II ISCED 2 (middle school)	7		8		9		
	ISCED 3C shorter than 2 years	**4**									
	ISCED 3C 2 years or more	**5**	Zasadnicze zawodowe	basic vocational	10		11		11		V
	ISCED 3A-B	**6**	Średnie zawodowe	secondary vocational	12		13		13		V
			Średnie ogólnokształcące	Secondary general	11		12		12		A
	ISCED 3 (without distinction A-B-C, 2y+)	**7**									
	ISCED 4C	**8**									
	ISCED 4A-B	**9**									
	ISCED 4 (without distinction A-B-C)	**10**	Policealne, pomaturalne, ale nie wyższe (ISCED 4)	post secondary, non-tertiary, ISCED 4	13		14		14		V

POLAND											
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed, born before 1952**	**School starting age**	**Total years of schooling when level is completed, born after 1951 and before 1986**	**School starting age**	**Total years of schooling when level is completed, born after 1985**	**School starting age**	**Vocational/ Academic**
	ISCED 5B	**11**									
	ISCED 5A, bachelor degree	**12**	Licencjat (ISCED 5A) (studia I stopnia)	BA, ISCED 5A (I degree)	14		15		15		
	ISCED 5A, master degree	**13**	Magisterium (ISCED 5A) (studia II stopnia)	MA, ISCED 5A (II degree)	16		17		17		
	ISCED 6	**14**	Doktorat, profesura (ISCED 6)	PhD, Professor, ISCED 6	20		21		21		

RUSSIAN FEDERATION*							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the following qualifications is the highest you have		Какой наивысший уровень образования Вы получили?	Which of the following qualifications is the highest you have			6
	No formal qualification or below ISCED 1	**1**	Без образования	Without education	6		
	ISCED 1	**2**	Не закончил школу (менее 9ти классов)	Doesn't graduated from secondary school (Less than 9 classes)	6		
	ISCED 2	**3**	9 классов средней школы	9 classes of secondary school	9		
	ISCED 3C shorter than 2 years	**4**					
	ISCED 3C 2 years or more	**5**					
	ISCED 3A-B	**6**					
	ISCED 3 (without distinction A-B-C, 2y+)	**7**	Средняя школа (10-11 классов)	secondary school (10-11 classes)	11	A	
	ISCED 4C	**8**					
	ISCED 4A-B	**9**					
	ISCED 4 (without distinction A-B-C)	**10**	профессиональное образование (например,	education in specialized school	11	V	
	ISCED 5B	**11**	Среднее профессиональное образование (например, техникум)	technical secondary school	12		
	ISCED 5A, bachelor degree	**12**	Незаконченное высшее образование, бакалавр	Incomplete higher education, bachelor	16		
	ISCED 5A, master degree	**13**	Высшее образование, магистр	Higher education, master's degree	18		
	ISCED 6	**14**	Ученая степень (кандидат, доктор наук) или два высших образования	Academic degree or two higher educations	21		

RUSSIAN FEDERATION*							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	Foreign qualification	**15**	Зарубежное образование	Foreign education			

* Please refer to the note regarding the Russian Federation in the Note to Readers section of this report.

SPAIN							
Int. Question No	**International English Version**	**Int. Value Code**		**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the following qualifications is the highest you have obtained?		De las titulaciones incluidas en esta ficha, ¿cuál es la más alta que ha obtenido?	Which of the following titles the highest you have finished?			5
	No formal qualification or below ISCED 1	**1**	Menos de 5 años de escolarización.	Infant education, "párvulos" school, nursery school and similar. We do not consider that there exists formal education at any level below primary education.			
	ISCED 1	**2**	Educación Primaria; 5 ó más años de escolarización; Educación General Básica (5 cursos); y similares.	Primary education, certificate of primary studies, Spanish languages for immigrants, EGB Basic General Education (years 1-5) and similar.	6		

SPAIN							
Int. Question No	**International English Version**	**Int. Value Code**		**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 2	**3**	Educación Secundaria Obligatoria; Certificado de Estudios Primarios;Educación General Básica (2ª etapa); Bachilleratos Elementales; Formación Profesional, programa de aprendizaje de tareas; Pruebas de acceso a ciclos formativos de Grado Medio; y similares.	Compulsory Secondary Education; Basic General Education (years 6-8), Elementary Baccalaureats; Vocational Education, programme for the learning of skills; Social guarantee programme in 1 year; Initial vocational quallification programme in 1 year; and similar.	10		
	ISCED 3C shorter than 2 years	**4**	Programa de Garantía Social; Programa de cualificación profesional inicial; y similares.	Professional technical studies for adults; occupational education, and similar	11	V	
	ISCED 3C 2 years or more	**5**	Grado Medio de Música y Danza; certificado de la Escuela Oficial de Idiomas; FPI; y similares.	Specific Vocational Education, Programme for Initial Vocational Qualification, in 2 years; middle level of Official Schools of Languages; tests for access to university for people over 25; former Vocational Education 1st level. And similar.	12	V	

SPAIN							
Int. Question No	**International English Version**	**Int. Value Code**		**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 3A-B	**6**	Bachillerato, antiguos Bachilleratos Superiores y cursos preuniversitarios; BUP,COU; Formación Profesional Específica, Artes Plásticas y Enseñanzas deportivas de grado medio;FPI; Oficialía; y similares.	Baccalaureate , former Higher Baccalaureates and pre-university courses. And similar.	12	A	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**					
	ISCED 4C	**8**					
	ISCED 4A-B	**9**	Pruebas de acceso a ciclos formativos de grado superior; y similares.	Tests to have access to Specific Vocational Education, higher level, and similar Tests to have access to Vocational Education, higher level, and similar	14	A	
	ISCED 4 (without distinction A-B-C)	**10**					
	ISCED 5B	**11**	Formación Profesional Específica, Artes plásticas y Enseñanzas deportivas de grado superior; FPII; Maestría industrial; y similares.	Specific Vocational Education, higher level; Higher Level of Music/Dance Conservatories; Higher level in Plastic Arts/Design/Sports Technician;and similar	14		

SPAIN							
Int. Question No	International English Version	Int. Value Code		English translation of the national version	Total years of schooling when level is completed	Vocational/ Academic	School starting age
	ISCED 5A, bachelor degree	**12**	Diplomatura; Ingeniería y Arquitectura técnica; licenciatura; estudios superiores de Artes Plásticas y Diseño; Estudios de Conservación y restauración; títulos de grado; y similares.	University Diploma; Tecnical Engineering and Architecture; University "Licenciatura", Higher Engineering and Architecture, titles of "Grado" and similar	15		
	ISCED 5A, master degree	**13**	Máster oficial;licenciatura; ingeniería superior y arquitectura; especialidades sanitarias de posgrado; y similares.	Master Degrees and postgraduate medical specializations. And similar.	17		
	ISCED 6	**14**	Doctorado.	Doctoral courses with a thesis, or equivalent requirements.	21		

SLOVAK REPUBLIC							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the qualifications on this card is the highest you have obtained?		Ktorý z nasledujúcich stupňov vzdelania je najvyšší, aký ste doteraz dosiahli?	What is the highest education you have completed?			6
	No formal qualification or below ISCED 1	**1**	Predškolská výchova				
	ISCED 1	**2**	Základná škola 1.-4. trieda	Primary school 1-4. years	4		
	ISCED 2	**3**	Zákl. škola 5.-9. trieda, 8 ročné gymnázium 1.-4. ročník, osobitná škola 5.-9. ročník	Lower secondarry 5.-9. years Lower secondary school for SEN 5-9. years	9		
	ISCED 3C shorter than 2 years	**4**	Stred. odborné školy, učilištia (kratšie ako 3 roky)	Secondary technical / vocational schools (les than 3 years)	11	V	
	ISCED 3C 2 years or more	**5**	Stred. odborné školy, učilištia (3 roky a viac)	Secondary technical / vocational schools (3 years or more)	12	V	
	ISCED 3A-B	**6**	Stredné školy s maturitou	Secondary schools with school leaving exam	13	A	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**	Pomaturitné vzdelávanie	Upper secondary school	14	V	
	ISCED 4C	**8**	Vyššie odborné školy, konzervatóriá 5.-6.ročník	Pre terciary school, Secondary art school 5-6 years	15	V	
	ISCED 4A-B	**9**					
	ISCED 4 (without distinction A-B-C)	**10**					
	ISCED 5B	**11**					
	ISCED 5A, bachelor degree	**12**	Vysokoškolské vzdelanie I. stupňa (Bakalárske štúdium, Bc.)	Bachelor degree, Gradual study	16		

SLOVAK REPUBLIC							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 5A, master degree	**13**	Vysokoškolské vzdelanie II. stupňa (napr. Mgr., Ing., MUDr., PhDr.)	Master degree	18		
	ISCED 6	**14**	Vysokoškolské vzdelanie III. stupňa (napr. PhD.)	PhD studies, Second stage of tertiary education	21		
	Foreign qualification	**15**					

SWEDEN							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed (for respondents who did not confirm validity of register data YRSQUAL was calculated on the basis of more detailed Swedish B_Q01aSE1 variable)**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the qualifications on this card is the highest you have obtained?		Vilken är den högsta utbildning som du har fullföljt?	What is the highest education you have completed?		Derived from register data	6
	No formal qualification or below ISCED 1	**1**	Grundskola, högst 5 år (folkskola/motsvarande) eller ingen utbildning alls	No formal education or education at primary level for 5 years or shorter (below ISCED 1)	6		
	ISCED 1	**2**	Grundskola, 6-8 år (folkskola/grundskola/motsvarande)	6 - 8 years of education at primary level (ISCED 1)	6		
	ISCED 2	**3**	Grundskola, enhetsskola eller realskola (9 - 10 år)	Completed compulsory school (9 - 10 years of education at lower secondary level) (ISCED 2)	9		
			Yrkesutbildning 2 månader - 1 år motsvarande heltid utöver folkskola/grundskola	Vocational education 2 months - 1 year corresponding to full time, based on elementary/comprehensive school (ISCED 2)	10 (9)*		

SWEDEN							
Int. Question No	International English Version	Int. Value Code	National Version	English translation of the national version	Total years of schooling when level is completed (for respondents who did not confirm validity of register data YRSQUAL was calculated on the basis of more detailed Swedish B_Q01aSE1 variable)	Vocational/ Academic	School starting age
			Grundskole-kompetens inom vuxen- utbildning, folkhögskola	Received certificate corresponding to education at lower secondary level from adult education or folk high schools (ISCED 2)	9		
			Flickskola	girls' school (ISCED 2)	10 (9)*		
	ISCED 3C shorter than 2 years	**4**	Gymnasieutbildning kortare än 2 år, Fackskola och yrkesutbildning kortare än 2 år	Education at upper secondary schools shorter than 2 years, vocational education shorter than 2 years (ISCED 3C < 2 years)	10		
	ISCED 3C 2 years or more	**5**	3-årigt gymnasium, även yrkes- utbildning 3 år	Education at upper secondary schools 2 years, vocational education 2 years (ISCED 3 2 years and more)	12		
	ISCED 3A-B	**6**	Gymnasieutbildning 2 år, Fackskola eller yrkes- utbildning 2 år	Education at upper secondary schools 3 years, vocational education 3 years (ISCED 3 2 years and more)	11		
	ISCED 3 (without distinction A-B-C, 2y+)	**7**					

SWEDEN							
Int. Question No	International English Version	Int. Value Code	National Version	English translation of the national version	Total years of schooling when level is completed (for respondents who did not confirm validity of register data YRSQUAL was calculated on the basis of more detailed Swedish B_Q01aSE1 variable)	Vocational/ Academic	School starting age
			Vuxen-utbildning mot-svarande gymnasium 2 - 3 år, även inom folkhögskola	Adult education on upper secondary level 2 - 3 years, also in folk high schools (ISCED 3 2 years and more)	12		
	ISCED 4C	**8**					
	ISCED 4A-B	**9**					
	ISCED 4 (without distinction A-B-C)	**10**	4-årigt gymnasium, basår eller påbyggnads-utbildning för vuxna utöver 3 års gymnasial utbildning	4 years education at upper secondary schools or supplementary education for adults (based on 3 years education at upper secondary school) (ISCED 4)	13		
			Högskole-/universitetsutbildning och annan eftergymnasial utbildning motsvarande heltidsstudier kortare än 2 år	Education at college/university or other post secondary education shorter than 2 years (ISCED 4)	13		

SWEDEN							
Int. Question No	**International English Version**	**Int. Value Code**	**National Version**	**English translation of the national version**	**Total years of schooling when level is completed (for respondents who did not confirm validity of register data YRSQUAL was calculated on the basis of more detailed Swedish B_Q01aSE1 variable)**	**Vocational/ Academic**	**School starting age**
	ISCED 5B	**11**	Högskole- /universitetsutbildning motsvarande 2 års heltidsstudier och annan eftergymnasial utbildning motsvarande 2 års heltidsstudier eller längre	Education at college/university 2 years or other post secondary education 2 years or longer (ISCED 5B)	14		
	ISCED 5A, bachelor degree	**12**	Högskole- /universitetsutbildning motsvarande 3 års heltidsstudier	Education at college/university 3 years (ISCED 5A)	15		
	ISCED 5A, master degree	**13**	Högskole- /universitetsutbildning motsvarande heltidsstudier i 4 år eller längre	Education at college/university 4 years or longer (ISCED 5A)	16		
	ISCED 6	**14**	Forskarutbildning (Fil lic eller Fil Dr)	Post graduate education (Licentiate of Ph or PhD degree) (ISCED 6)	20		
		* for respondents who confirmed validity of register data YRSQUAL was calculated on the basis of international B_Q01a variable. Therefore, Swedish B_Q01aSE1 categories belonging to international value code 3 could not be distinguished. They were all assigned value 9 as minority of respondents belong to Yrkesutbildning and Flickskola					

UNITED STATES							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
B_Q01a	Which of the qualifications on this card is the highest you have obtained?		Looking at this card, what is the highest level of education you have completed?				5
	No formal qualification or below ISCED 1	**1**	Pre-primary or no schooling				
	ISCED 1	**2**	Grades 1-6		6		
	ISCED 2	**3**	Grades 7-9		9		
	ISCED 3C shorter than 2 years	**4**					
	ISCED 3C 2 years or more	**5**					
	ISCED 3A-B	**6**					
	ISCED 3 (without distinction A-B-C, 2y+)	**7**	High school diploma		12	A	
	ISCED 3 (without distinction A-B-C, 2y+)	**7**	Pre-associate education. Attended trade school, college, or university; no certificate or degree received.		NA	A	
	ISCED 4C	**8**					
	ISCED 4B	**9**					
	ISCED 4A	**9**	A certificate from a college or trade school for completion of a program prior to the associate/bachelor’s degree.		13	V	
	ISCED 4 (without distinction A-B-C)	**10**					
	ISCED 5B, associate degree	**11**	Associate degree		14		
	ISCED 5A, bachelor degree	**12**	Bachelor's degree (e.g. BA, AB, BS)		16		

UNITED STATES							
Int. Question No	**International English Version**	**Int. Value Code**	**National version (text)**	**English translation of the national version**	**Total years of schooling when level is completed**	**Vocational/ Academic**	**School starting age**
	ISCED 5A, master degree	**13**	Master's degree (e.g. MA, MS, Meng, MEd, MSW, MBA)		18		
	ISCED 5A, master degree	**13**	Professional degree (e.g. MD, DDS, DVM, LLB, JD)		19		
	ISCED 6	**14**	Doctorate degree (e.g. PhD, EdD)		21		

Appendix 6: PIAAC Consortium – Staff, Expert Group and National Project Manager Listings

PIACC Consortium

Educational Testing Service (ETS) – Overall Management, Test Development, Psychometrics, Analysis and Data Products

Irwin Kirsch (International Project Director)

Claudia Tamassia (International Project Manager)

Kentaro Yamamoto (Director, Psychometrics and Analysis)

Matthias von Davier (Co-Director, Psychometrics and Analysis)

Marylou Lennon (Test Development, Literacy and PSTRE)

John P. Sabatini (Test Development, Reading Components)

Kelly M. Bruce (Test Development, Reading Components)

Eugenio Gonzalez (Training and Technical Report)

Michael Wagner (Director, Platform Development)

Larry Hanover (Editorial Support)

Judy Mendez (Project Support)

Lisa Hemat (Project Support)

Jason Bonthron (Platform Development)

Mike Ecker (Platform Development)

Ramin Hemat (Platform Development)

Tom Florek (Platform Development)

Debbie Pisacreta (Platform Development)

Janet Stumper (Platform Development)

John Barone (Director, Data Analysis and Database Preparation)

Scott Davis (Data Analysis)

Justin Herbert (Data Analysis)

Steven Holtzman (Data Analysis)

Laura Jerry (Data Analysis)

Mathew Kandathil (Data Analysis Leader)

Debra Kline (Data Management)

Nan Kong (Data Analysis)

Phillip Leung (Data Analysis Leader)

Chen Li (Data Analysis)

Mei-Jang Lin (Data Analysis)

Michael Narcowich (Data Analysis)

Alfred Rogers (Data Analysis Leader)

Jonathan Steinberg (Data Analysis)

Joan Stoeckel (Data Analysis and Data Management)

Ruopei Sun (Data Analysis)

Minhwei Wang (Data Analysis Leader)

Kei Sing Wong (Data Analysis)

Lingjun Wong (Data Analysis)

Jeffrey Wright (Data Analysis)

Fred Yan (Data Analysis)

Ningshan Zhang (Data Analysis)

Danielle Baum (Consultant, Paper Booklets)

Juliette Mendelovits (Consultant, Literacy Test Development, ACER)

Dara Searle (Consultant, Literacy Test Development, ACER)

GESIS – *Development of the Job Requirement Approach Module and Background Questionnaire*

Beatrice Rammstedt (Lead)

Dorothée Behr

Susanne Helmschrott

Silke Martin

Natascha Massing

Anouk Zabal

Deutsches Institut für Internationale Pädagogische Forschung (DIPF) – *Development of the PIAAC Test Delivery Platform*

Ingo Barkow (International IT Support)

Robert Baumann (Software Development)

Simon Brüchner (Software Development)

Mahtab Dalir (Software Development)

Gabriele Gissler (Item Development)

Frank Goldhammer (Test Development, Deputy Project Co-Director)

Roland Johannes (Software Development)

Elham Müller (Software Development)

Jean-Paul Reeff (International Consultant)

Marc Rittberger (Director)

Heiko Rölke (Project Co-Director)

Alexander During (Software Development)

Maya Schnitzler (Software Development)

Felix Toth (Software Development)

Britta Upsing (Project Coordinator)

cApStAn – *Linguistic Quality Control*

Steve Dept (Verification Operations)

Andrea Ferrari (Verification Methodology and Management)

Laura Wäyrynen (Verification Methodology and Management)

Elica Krajčeva (Verification Management)

Raphaël Choppinet (Verification Management)

Shinoh Lee (Verification Management)

Irene Liberati (Verification Management)

Research Centre for Education and the Labour Market (ROA), Maastricht University – *Development of the Job Requirement Approach Module and Background Questionnaire*

Rolf van der Velden (Coordinator, Development Background Questionnaire)

Jim Allen (Development Background Questionnaire)

Martin Humburg (Development Background Questionnaire)

International Association for the Evaluation of Educational Achievement (IEA) – *Data Cleaning and Database Preparation*

Alena Becker (Data Processing and National Adaptations)

Christine Busch (Meta-data and Processing)

Ralph Carstens (Lead International Data Management and Analysis Support/Training)

Mark Cockle (Quality Control and Manuals)

Tim Daniel (Co-Lead International Data Management)

Bastian Deppe (Software Testing and Data Cleaning)

Limiao Duan (Processing Systems Development)

Daniela Tranziska (Processing Systems Development)

Christian Harries (Software Development)

Pamela Inostroza (Processing Systems Development)

Matthias Jenzen (Software Development)

Maike Junod (Software Development)

Alexander Konn (Processing Systems Development)

Kamil Kowolik (Data Processing and National Adaptations)

Alexander Lebedev (Software Testing)

Sebastian Meyer (Data Processing and National Adaptations)

Pia Möbus (Software Testing and Data Cleaning)

Jirka Neumann (Data Processing and National Adaptations)

Brice Nzuakue Diogni (Software Testing)

Dirk Oehler (Quality Control and Processing Systems)

Martin Olszewski (Processing Systems Testing)

Daniel Radtke (Data Processing and National Adaptations)

Frank Wohnfurter (Software Development)

Westat – *Sample Design and Selection, Weighting, Survey Operations, and Quality Control*

Leyla Mohadjer (Director, Sampling Activities)

Pat Montalvan (Director, Survey Operations)

Tom Krenzke (Manager, Sampling Activities)

Michael Lemay (Manager, Survey Operations)

Wendy Van de Kerckhove (Senior Leader, Sampling Activities)

Valerie Hsu (Leader, Sampling Activities)

Laura Alvarez-Rojas (Senior Survey Statistician)

Lillian Diaz-Hoffmann (Survey Operations Material Development and Training)

Sylvia Dohrmann (Senior Survey Statistician)

Jarrod Grebing (Survey Operations Training)

Hongsheng Hao (Senior Survey Statistician)

Wen-Chau Haung (Senior Systems Analyst)

Michael Jones (Senior Survey Statistician)

Robin Jones (Senior Systems Analyst)

Jane Li (Senior Survey Statistician)

Lin Li (Senior Survey Statistician)

Yuki Nakamoto (Senior Systems Analyst)

Margo Tercy (Project Support)

Klaus Teuter (Senior Systems Analyst)

Chao Zhou (Survey Statistician)

Public Research Center Henri Tudor – *Development of the Computer-Based Platform for the Background Questionnaire*

Thibaud Latour (Scientific Unit Leader, Project Coordination)

Isabelle Jars (Project Management)

Raynald Jadoul (Software Architecture and Staff Coordination)

Patrick Plichart (Platform Architecture)

Vincent Porro (Lead Designer and Development)

Lionel Lecaque (Platform Integration)

Jérôme Bogaerts (Lead Developer)

Joël Billard (Questionnaire Development)

Damien Arcani (Contents Designer)

Somsack Sipasseuth (Workflow Development)

Primaël Lorbat (Multilingual Framework Development)

Younes Djaghloul (Multilingual Framework Development)

Igor Ribassin (Virtual Machine Integration)

Pierre Goulaieff (Communication)

Expert Groups

PIAAC Literacy Expert Group

Stan Jones (Chair), Canada

Egil Gabrielsen, Center for Reading Research, University of Stavanger, Norway

Jan Hagston, Australia

Pirjo Linnakylä, University of Jyväskylä, Finland

Hakima Megherbi, University of Paris, France

John Sabatini, Educational Testing Service, United States of America

Monika Tröster, German Institute of Adult Education, Germany

Eduardo Vidal-Abarca, Department of Psychology, Universidad de Valencia, Spain

PIAAC Numeracy Expert Group (Test Development, Numeracy)

Iddo Gal (Chair), University of Haifa, Israel

Silvia Alatorre, National Pedagogical University, Mexico

Sean Close, St. Patrick's College, Ireland

Jeff Evans, Middlesex University, United Kingdom

Lene Johansen, Aalborg University, Denmark

Terry Maguire, Institute of Technology Tallaght-Dublin, Ireland

Myrna Manly, United States of America

Dave Tout, Australian Council for Educational Research, Australia

PIAAC Problem Solving in Technology-Rich Environments Expert Group

Jean-François Rouet (Chair), CNRS and University of Poitiers, France

Mirelle Bétrancourt, University of Geneva, Switzerland

M. Anne Britt, Northern Illinois University, United States of America

Dr. Rainer Bromme, University of Muenster, Germany

Arthur C. Graesser, University of Memphis, United States of America

Jonna M. Kulikowich, Pennsylvania State University, United States of America

Donald J. Leu, University of Connecticut, United States of America

Naoki Ueno, Musashi Institute of Technology, Japan

Herre van Oostendorp, Utrecht University, Netherlands

PIAAC Questionnaire Expert Group

Ken Mayhew (Chair), Pembroke College, Oxford, the United Kingdom

Patrice de Broucker, Statistics Canada, Canada

Enrique Fernandez, European Foundation for the Improvement of Living and Working Conditions in Dublin, Ireland

Masako Kurosawa, National Graduate Institute for Policy Studies, Japan

Scott Murray, DataAngel Policy Research Incorporated, Canada

Jürgen Schupp, German Institute for Economic Research DIW in Berlin, Germany

Tom W. Smith, University of Chicago, United States of America

Kea Tijdens, University of Amsterdam, the Netherlands

Robert Willis, Michigan, United States of America

PIAAC Technical Advisory Group

Cees A. W. Glas (Chair), University of Twente, the Netherlands

Thomas Amosse, France

Roel Bosker, University of Groningen, the Netherlands

Henry Braun, Boston College, United States of America

Lars Lyberg, Stockholm University, Sweden

Robert Mislevy, University of Maryland, United States of America

Christian Monseur, University of Liège, Belgium

Irini Moustaki, London School of Economics, the United Kingdom

National Project Managers

Australia: Wendy Ozols

Austria: Markus Bönisch

Belgium: Inge De Meyer

Canada: Sylvie Grenier

Cyprus: Athena Michaelidou

Czech Republic: Jana Strakova

Denmark: Anders Rosdahl

Estonia: Aune Valk

Finland: Antero Malin

France: Nicolas Jonas

Germany: Beatrice Rammstedt

Ireland: Donal Kelly

Italy: Gabriella Di Francesco

Japan: Atsushi Kogirima

Korea: Eon Lim

Netherlands: Willem Houtkoop

Norway: Birgit Bjørkeng

Poland: Jan Burski

Slovak Republic: Adriana Mesarosova

Spain: Luis Sanz and Ines Sancha

Sweden: Ann-Charlott Larsson

United Kingdom: Julie Sewell and Rebecca Wheater

United States: Eugene Owen

Appendix 7: Data Adjudication in PIAAC

PIAAC Consortium and William Thorn

Section A7-1. Data adjudication – content, process and outcome

This section describes the content and process for the evaluation of quality – known as adjudication – of the data collected by participating countries, and provides a brief summary of the outcome of the process. The objective of the data adjudication process was to arrive at a judgment regarding the global quality of the data from PIAAC for each participating country and to determine, if necessary, any limitations that should be applied to the public dissemination and use of these data.

The PIAAC Technical Standards and Guidelines (TSG)[1] established requirements relating to the quality of PIAAC survey data with respect to representation of the target population and data comparability across countries, and provided standard procedures for quality assurance. Throughout the survey process, the Consortium conducted continuous quality monitoring activities aimed at limiting the magnitude of quality variation among countries. Communication between the country and its assigned Consortium contacts for sampling, operations and other components of the survey was critical to understanding various aspects of country samples and for assessing the quality and comparability of PIAAC data nationally and across countries. Communication allowed the Consortium to recommend ways to improve the quality of the country samples at the same time as minimizing the quality variation among countries.

The quality control (QC) process collected information regarding the country status following the TSG. The National Survey Design and Planning Report was the initial tool for collecting information from the countries about country-specific approach to maintaining compliance with the TSG for the total survey process. The implementation of those planned processes was monitored closely. For example, operations were monitored through conference calls on a regular basis and reports provided from the country relating to response rates and validation. Also, for sampling, the primary vehicles for the communication were the QC Sample Selection and Sample Monitoring forms. Real-time monitoring of all aspects of sampling was critical in allowing the Consortium to uncover problems with sampling and for the countries to incorporate changes if necessary.

As emphasized above, compliance with the TSG was an important component in the assessment of national data. However, in the adjudication process, a wider definition of quality was used – that of

[1] The December 2010 version of the TSG can be accessed from the following link: http://www.oecd.org/site/piaac/PIAAC-NPM(2010_12)PIAAC_Technical_Standards_and_Guidelines.pdf (accessed 24 September 2013).

"fitness for use." In other words, the goal was to go beyond compliance to assess whether the data produced were of a sufficient quality in terms of their intended uses or applications. In assessing overall quality level, the focus was on four key areas:

- Sampling
- Coverage and nonresponse bias
- Data collection
- Instrumentation

The core element of the adjudication process was an assessment of the quality of data in each of the domains identified above in terms of performance against a set of quality indicators. These indicators are listed in Table A7-1 below and described in detail in Section A7-2. These indicators reflect the major requirements of the TSG in the domains concerned and help to assess the variation in quality when attempting to compare estimates across countries.

Table A7-1: Quality domains and associated indicators

Domain	Indicators[2]
1- Sampling	1.A Sampling plan[3] 1.B Sample selection (home office)[2] 1.C Sample selection (field) 1.D Sample weighting 1.E Sampling error (DEF)
2 - Coverage and nonresponse bias (NRBA)	2.A Population coverage (frame) 2.B Population coverage (field) 2.C Weighted response rate, and coverage rate 2.D NRBA (Basic) 2.E NRBA (extended)
3 - Data collection	3.A Field validation/rechecks[4] 3.B Staffing, training, management / monitoring
4 - Instrumentation	4.A Assessment data 4.B Background questionnaire data 4.C Translation 4.D Coding and scoring 4.E Item nonresponse

In each of the four domains, the Consortium made an assessment of the level of performance of countries, first, at the level of each of the individual indicators and, second, at the level of the domain as a whole (see Table 2). A three-category assessment schema was used to summarize the assessments in respect to each indicator and, globally for each domain – “passed” (i.e., relevant requirements completely met), “caution” (i.e., relevant requirements met to a reasonable extent) and “failed” (i.e., relevant requirements generally not met). Explanations of what the assessment categories mean in relation to the quality indicators are provided in Section A7-2. At the level of individual indicators, the assessment was based on compliance with relevant standards, the information provided by countries as part of the quality control process, and the analysis of the response data from the Main Study. At the level of the domain, the assessment was based on consideration of performance in relation to the relevant indicators and their interrelationships. For example, evidence of a high level of undercoverage bias could be judged to be a serious problem for quality even if response rates were high and nonresponse bias low.

[2] Indicator codes as in Annex 1.

[3] The goal of the Consortium was to have the sampling plan and sample selection verified for all countries before they went to the field. The schedule was set up so countries had enough time to incorporate corrections to their sampling steps before data collection. However, a number of countries had major delays in submitting their forms and thus there was no chance for correcting errors or improving upon deficiencies, if any, in these samples.

[4] Data collection validation (rechecks) is critical to data validity; it is the most important quality control feature of household data collection. However, because this is the first cycle of PIAAC, it was understood if a country had not fully met the standards surrounding this activity. However, serious consideration will be given to raising the importance of this adjudication feature for the next cycle of PIAAC.

Table A7-2: Levels of quality assessment

Domain	Assessment against Indicators	Overall Assessment
Sampling	1.A (pass, caution, fail) 1.B (pass, caution, fail) 1.C (pass, caution, fail) 1.D (pass, caution, fail) 1.E (pass, caution, fail)	Pass, caution, fail
Coverage and nonresponse bias	2.A (pass, caution, fail) ...	Pass, caution, fail
Data collection	3.A (pass, caution, fail) ...	Pass, caution, fail
Instrumentation	4.A (pass, caution, fail) ...	Pass, caution, fail

The Consortium summarized the outcome of the assessment for each indicator and domain in an initial report that presented the results to the PIAAC Technical Advisory Group (TAG) at its meeting in December 2012. Countries were provided with the initial report soon after it was circulated to the TAG. The TAG reviewed the results of the quality assessment (and any country responses) and provided a report to the OECD Secretariat and to the BPC containing recommendations regarding the presentation and use of data for each country. The adjudication process was finalized after countries performed the required analyses based on the proficiency estimates that became available to countries prior to the 16-18 January 2013 workshop in Paris.

A decision was made to recommend that some conditions be placed on the release of a country's data if it received a "fail" grade in one or more domains. The conditions could range from placing results from the country concerned "under the line" in tables accompanied by an appropriate annotation when reporting results to suppressing data in some tables or, at the extreme, not releasing a country's data as part of the PIAAC international dataset. Similarly, receipt of a "caution" for two or more domains could lead to a recommendation that conditions be placed on release of a country's data.

The content of any recommendations made regarding the conditions applying to the release of a country's data reflected, in addition to the principles articulated in the TSG regarding response rates (Standard 4.7.4 and associated guidelines) described in Table 5 below, the extent and nature of the problems concerning data quality.

From the point of view of providing a secure basis for making inferences regarding the target population, some indicators are more important than others. For example, as discussed in Chapter 14, probability sampling is a necessary condition for a representative sample. Failure to provide evidence that sample selection both at the design stage and in the field resulted in a probability sample would represent a more serious concern than a failure to follow the standards relating to the training of interviewers and would lead, other things being equal, to more stringent conditions being placed on data release. Also, for example, as given in Chapter 16, design effects (DEFFs) are an example of one of the more visible indicators of quality variation among countries. DEFFs are a measurable summary of quality and take into account the impact on sampling error due to clustering, stratification, unequal probabilities of selection, weight adjustments (Chapter 15) and multiple imputation. Design effects were estimated prior to sample selection, and for countries with relatively high design effects it was recommended to attempt to improve the stratification in their designs by finding good correlates with the PIAAC outcomes, and to revisit the clustering in their sample designs.

As discussed in the TSG, given the relationships between bias and undercoverage and response rates, countries must keep the exclusion rates low and implement procedures to reduce the potential for nonresponse bias and attain high response rates. There were several ways to reduce the potential for nonresponse bias. First and foremost was to plan and implement field procedures that obtain a high level of cooperation. Response rate was a valuable data quality component of the analysis of nonresponse bias, which was an important input to the data adjudication process. As explained in Section A7-2 below, two types of nonresponse bias analysis (NRBA) were required from countries. All countries were to complete a basic NRBA designed to provide evidence for the selection of variables to be used in nonresponse weighting adjustments. Countries with overall response rates of less than 70 percent were also required to complete an extended NRBA designed to evaluate the impact of the weighting adjustments implemented on the proficiency estimates. Chapter 16 includes a description of the basic and extended NRBA, and the outcome of the analysis is included in the country reports in Section A7-3.

Table 3 presents the PIAAC Data Quality Evaluation results for all quality indicators. Section A7-3 includes each country’s adjudication report.

Table A7-3: PIAAC data quality evaluation summary table[5]

Country	Sampling	Coverage and Nonresponse Bias[6]	Data Collection	Instrumentation
Australia	Caution-Quality partially known, due to confidentiality restrictions	Pass	Pass	Pass
Austria	Pass	Caution-Bias low	Pass	Pass
Canada	Pass	Caution-Bias minimal	Pass	Pass
Cyprus[7]	Pass	Pass	Pass	Pass
Czech Republic	Pass	Caution-Bias low	Pass	Pass
Denmark	Pass	Caution-Bias low	Pass	Pass
England (UK)	Pass	Caution-Bias low	Caution-Partial Compliance	Pass
Estonia	Pass	Caution-Bias low	Pass	Pass
Finland	Pass	Caution-Bias minimal	Caution-Partial Compliance	Pass
Flanders (Belgium)	Pass	Caution-Bias low	Pass	Pass
France	Pass	Caution-Bias low	Caution-Partial Compliance	Pass
Germany	Caution-Probabilities of selection derived from simulation	Caution-Bias low	Pass	Pass
Ireland	Pass	Pass	Pass	Pass
Italy	Pass	Caution-Bias low	Pass	Pass

[5] This table represents summarized information that is extracted from Tables A7-4, A7-5 and A7-6 at the end of this Annex.
[6] The ratings provided in this column are based on sample coverage, response rate, and the outcome of NRBA. The analysis showed that nonresponse adjustment weighting was effective in reducing the potential for bias in all countries. However, there is still a potential for either minimal or low level of bias in the outcome statistics for countries with response rates lower than 70%. The analysis concluded that there was not enough evidence showing any moderate or high level of bias, based on assumptions made about the proficiency scores of nonrespondents. Therefore, data users need to be cautioned when interpreting the results of the analysis for countries with very low response rates because different assumptions could lead into different results.
[7] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table A7-3 (cont.): PIAAC data quality evaluation summary table

Country	Sampling	Coverage and Nonresponse Bias	Data Collection	Instrumenta-tion
Japan	Caution-Approved deviation from standards	Caution-Bias low	Pass	Pass
Korea	Pass	Pass	Pass	Pass
Netherlands	Pass	Caution-Bias low	Pass	Pass
Northern Ireland (UK)	Pass	Caution-Bias low	Pass	Pass
Norway	Pass	Caution-Bias low	Pass	Pass
Poland	Pass	Caution-Bias low	Caution-Partial Compliance	Pass
Russian Federation[8]	Caution – Noncompliance	Caution-Bias level unknown level[9]	Fail	Caution[10]
Slovak Republic	Pass	Caution-Bias low	Pass	Pass
Spain	Pass	Caution-Bias low	Pass	Pass
Sweden	Pass	Caution-Bias low	Pass	Pass
United States	Pass	Pass	Pass	Pass

[8] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.
[9] Bias level unknown due to incomplete nonresponse bias analyses.
[10] See "Data Adjudication Summary" section in the Russian Federation Adjudication Report for details.

Section A7-2: Data quality – indicators used for adjudication

1. Sampling

1.A Sampling plan

- A complete sampling plan was provided.
- The country responded to feedback from the Consortium.

Rating: “Pass” = requirements fully met; “caution” = plan provided but only limited response to suggestions; “fail” = no plan provided, plan provided but country did not respond to feedback.

1.B Sample selection: Home office

- Complete QC sample selection forms were provided prior to data collection.
- Each person in the PIAAC target population had a nonzero and known (calculable) probability of selection resulting from the application of established and professionally recognized principles of scientific sampling.
- No substitution of sampling units.

Rating: “Pass” = requirements fully met; “caution” = evidence that sample selection process was not based on probability principles, but that effects were not significant; “fail” = no information provided, evidence that sample selection process was not based on probability principles and that effects were potentially significant.

1.C Sample selection: In field

- Persons were selected from within households using a fully enumerated grid of household members.
- No more than two persons were selected in a household, and fewer than 10% of households had two persons selected.
- Each person in the PIAAC target population had a nonzero and known (calculable) probability of selection resulting from the application of established and professionally recognized principles of scientific sampling.
- No substitution of sampling units.

Rating: “Pass” = requirements fully met; “caution” = only partial information provided or evidence that sample selection process was not based on probability principles, but that effects are not significant; “fail” = no information or insufficient provided, evidence that sample selection process was not based on probability principles and that effects were potentially significant.

1.D Sample weighting

- The country fully completed and returned the applicable QC weighting forms.

- Persons who did not complete the survey for a literacy-related reason (e.g., language barrier) were excluded from the adjustment for nonliteracy-related nonresponse. Literacy-related nonrespondents (LRNR) at the screener stage or without age and gender collected were represented by BQ LRNR with age and gender collected and assessment LRNR. The BQ LRNR with age and gender collected has final weights and was included in the benchmarking adjustment with the BQ respondents.
- At a minimum, weights were benchmarked to control totals for age and gender.
- Control totals were from a survey of higher quality than PIAAC and match the concepts and definitions in PIAAC.
- Between 15 and 80 replicate weights were created using one of the following methods: delete-one jackknife, paired jackknife, balanced repeated replication, or Fay's method.
- All weight adjustments conducted for the full sample were conducted on each replicate weight to capture the variation created, or reduced, by the weight adjustments.

Rating: "Pass" = requirements fully met; "caution" = requirement generally met; "fail" = requirements met to a very limited extent or not at all.

1.E Sampling error

- The design effect, as a result of clustering, differential sampling rates and weighting adjustments, is at an adequate level (less than 2.5) for proficiency measures. Two statistics are computed: 1) the unequal weighting effect, resulting from variable sampling weights, and 2) effective sample size, as the ratio of the final sample size and the design effect computed using the first plausible value for the literacy component.

Rating: "Pass" = requirements fully met; "caution" = requirement generally met; "fail" = requirements met to a very limited extent or not at all.

2. Coverage and nonresponse bias

2.A Population coverage: Frame

- The estimated percentage of the target population excluded from the frame

Rating: "Pass" = exclusions ≤ 5%; "caution" = 5% < exclusions ≤ 8%; "fail" = exclusions > 8%

2.B Population coverage: Data collection

- The weighted percentage of cases excluded because they are inaccessible. Rating: not applicable. This is provided as an information item.

2.C Weighted response rate

- The value of the overall design weighted response rate.

Rating: Pass" = response rate ≥70%; "caution" =50% ≤ response rate < 70%; "fail" = response rate < 50%

- The value of the overall design weighted coverage rate.

Rating: not applicable. This is provided as an indication of the overall coverage of the target population.

2.D NRBA: Basic

- The country performed all required basic NRBA analyses and returned the basic NRBA report.
- Variables related to age, gender, education, employment and region were analyzed.
- Characteristics showing bias were used in weighting adjustments or justification was provided for not including the variable in weighting.

Rating: "Pass" = requirements fully met; "caution" = requirement generally met; "fail" = requirements met to a very limited extent or not at all.

2.E NRBA: Extended 1-5 (only required if the overall weighted response rate was < 70%)

- The country completed the required analyses and returned the extended NRBA report.
- No evidence of significant, substantial undercoverage or nonresponse bias.

Rating: "Pass" = required analysis undertaken. No evidence of significant or substantial undercoverage or nonresponse bias; "caution" = required analysis undertaken. Evidence of a moderate level of undercoverage or nonresponse bias; "fail" = required analysis either not undertaken or undertaken to a limited extent. Evidence of a high level of undercoverage or nonresponse bias.

3. Data collection

3A. Validation/rechecks

- Overwhelming majority of validation cases were selected randomly.
- Close to 10% of each interviewer's cases were validated.
- Cases selected for validation included completes, refusals, noncontacts and ineligibles.

Rating: "Pass" = evidence provided that demonstrates that requirements were fully met; "caution" = evidence provided that demonstrates that requirements were generally met; "fail" = no information provided or available evidence indicates that requirements were not met or met only to a very limited extent.

3B. Data collection (staffing, training, management/monitoring)

- Sufficient and qualified staff were hired to conduct data collection (i.e., obtain required number of completes and acceptable response rates within the study timeframe).
- Interviewer training was conducted using adapted Consortium training scripts.
- Depending on experience, interviewers were offered at least 20-30 hours of in-person training.

- Interviewer training consisted of at least 10 hours covering BQ and direct assessment administration and four hours on gaining respondent cooperation.
- Field supervisors were responsible for no more than 30 interviewers.
- Meetings between interviewers and supervisors to manage and monitor field work were held at least every other week.

Rating: "Pass" = evidence provided that demonstrates that requirements were fully met; "caution" = evidence provided that demonstrates that requirements were generally met; "fail" = no information provided or available evidence indicates that requirements were not met or met only to a very limited extent.

4. Instrumentation

4.A Cognitive assessment

- Literacy, numeracy and problem-solving scales are reliable, valid and comparable.

Rating: "Pass" = significant deviations from international item characteristic curves (ICCs) observed in only a small number of cases; "caution" = significant deviations from international ICCs observed in some cases; "fail" = significant deviations from international ICCs observed in a large number of some cases.

4.B BQ

- BQ items and indices are reliable, valid and comparable.

Rating: "Pass" = data quality high (e.g., low levels of item nonresponse for key variables, scales reliable); "caution" = data quality moderate; "fail" = data quality low.

4.C Translation

- Translation conducted by two independent translators, followed by reconciliation by a third translator.
- Full verification undertaken before the Field Test, partial verification of any revisions undertaken before the Main Study.
- All BQ adaptations approved.

Rating: "Pass" = requirements fully met; "caution" = requirements generally met; "fail" = requirements met to a very limited extent or not at all.

4.D Coding and scoring

- Rates of agreement between countries of scoring of anchor booklets (literacy, numeracy).
- Level of scoring reliability within countries.
- Countries provided a description of their coding system and coding quality control procedures.

Rating: “Pass” = Evidence that the required scoring reliability studies were conducted correctly. Interrater reliability between and within countries was within expected bounds. Required information on coding provided; “caution” = evidence that the required scoring reliability studies may not have been conducted correctly. Interrater reliability between and/or within countries was outside expected bounds. Not all information on coding provided; “fail” = evidence that the required scoring reliability studies were not been conducted correctly. Interrater reliability between and/or within countries was well outside expected bounds. Required information on coding either not provided or only limited information provided.

4.E Item nonresponse

- Number of BQ items for which response rate is less than 85%.
- Item nonresponse bias analysis conducted for all BQ items with response rates below 85%.

Rating: “Pass” = requirements fully met; “caution” = requirement generally met; “fail” = requirements met to a very limited extent or not at all.

Section A7-3: Adjudication reports

Australia

Sampling

To the best of the Consortium's knowledge, Australia followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting, with one minor deviation (noted below under sample weighting). However, Australia was unable to provide all QC information because of confidentiality restrictions, so the Consortium cannot fully verify its compliance.

- Sampling plan: No issues
- Sample selection
  - Home office: For confidentiality reasons, Australia was unable to provide the Consortium with selection probabilities and could not report on most of the information in the QC sample selection forms. Therefore, the Consortium is unable to verify whether the sample adheres to the TSG.
  - In field: See above
- Sample weighting: For confidentiality reasons, Australia was unable to provide some of the information in the standard weighting QC forms. However, the Consortium corresponded with Australia to verify whether the main weighting standards were met. Australia performed person-level nonresponse adjustments and benchmarking to adjust for undercoverage and nonresponse at the household and person level, rather than doing separate adjustments at the household and person level according to the standard weighting procedures in the PIAAC Weighting and Variance Estimation Plan. Its procedure included a separate adjustment for literacy-related nonrespondents, as required by the TSG. The replicates were adjusted at each calibration stage but were not adjusted for nonresponse, which is in violation of Standard 14.11. However, per Australia, "Since the [nonresponse adjustment] factors are derived at such a broad level, they would vary very little if derived separately for each replicate group. Whilst this theoretically may result in variances being understated, in practice the magnitude of the impact is unlikely to be discernible."
- Sampling error: Australia's design effect due to unequal weights is 1.60 for a sample size of 7,428 adults ages 16-65. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 3,061. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (2.39). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. Australia produces both National and State level estimates for PIAAC so there are different probabilities of selection across the States/Territories. Since this survey design feature increases the design effect for the National estimates, Australia

increased its sample size to account for it. Further variation in the weights was added through within-household sampling, nonresponse and calibration adjustments, but if a weight was lower than 50% or higher than 300% of the initial weight after adjustments and benchmarking, benchmark classes were collapsed to reduce the weight fluctuation.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 3.3% (persons living in very remote areas; persons living in discrete indigenous communities; persons residing in non-institutional special dwellings; non-Australian diplomats, diplomatic staff and members of their household; non-Australian defense forces and their dependants).
  - Data collection: Not applicable
- Weighted response rate: 71%
- Nonresponse bias analysis
  - Basic: Instead of the standard analyses required by the Consortium, Australia performed a coverage analysis and calculated BQ response rates by subgroup. The Consortium agreed that the coverage analysis could serve as a substitute for the chi-square analysis. It encouraged Australia to perform a multivariate analysis (e.g., logistic regression). However, Australia explained that its coverage analysis is iterative—the potential bias after standard calibration is looked at first (by comparing weighted estimates to external totals) and then the weights are calibrated further if necessary. This is done in a way that would serve a similar purpose as a multivariate analysis.

    Australia evaluated nonresponse by region, but could not share the results because of confidentiality reasons. As well, an under-representation was found of males, younger age groups, less educated, and not employed. Gender, age, education, Labor force status and region were all used in weighting adjustments. No other variables were analyzed for nonresponse bias.
  - Extended (preliminary): Not required

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Australia generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 10.9.3 on fieldwork validation and Guidelines 8.1.1B and 8.1.2A on management of field staff.

However, Australia met a reduced requirement on interviewer training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. Australian interviewers were provided with 20 to 28 hours of in-person training.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Australia followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[11] prior to the Field Test and a partial verification[12] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Australia followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 98.3%
    - Literacy Items: 98.8%
    - Numeracy Items: 96.3%
- Scoring reliability of paper-based national booklets
    - Core items: 99.7%
    - Literacy Items: 98.1%
    - Numeracy Items: 99.2%

Assessment data

Overall, 96.5% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Australia, 78.0% of the respondents who completed the BQ took the computer-based cognitive assessment, while 19.7% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Australia, 14.4% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 2.8% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

[11] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[12] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Australia followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Australia. If a respondent started the interview, the likelihood that she/he provided data is at a level above 98% with practically only one exception: Income related questions. In Australia, 93.2% of respondents provided yearly income reported in either direct amount or categories.

If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Australia, we observed 1.9% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Australia, these percentages were 10.2% for Literacy and 7.4% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Australia, the percentage of nonresponse for Literacy was 5.9%, for Numeracy it was 4.6%, and for PSTRE it was 0.2%.

Austria

Sampling

Austria followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. All QC materials related to sampling plan, sample selection, and sample weighting were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
  - Home office: No issues
  - In field: Not applicable
- Sample weighting: Austria followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights.
- Sampling error: Austria's design effect due to unequal weights is 1.09 for a sample size of 5,130. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 3,561. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (1.41). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 0.6% (undocumented immigrants).
  - Data collection: The weighted percentage of cases excluded because they were inaccessible was 0.8%.
- Weighted response rate: 53%
- Nonresponse bias analysis
  - Basic: Austria performed all required analyses. The basic analysis showed significantly low response rates for low educated, non-Austrian, and people living in Styria and Vienna, based on registry information. Age, gender, province, urbanization, education, and nationality were used in weighting adjustments.
  - Extended: Austria performed all required analyses except the analysis for non-interview report form. The extended analysis showed that bias was reduced through the weighting adjustments.
    - Analysis 1 – Comparisons of estimates before and after weighting: Bias in age, education, nationality, urbanization, and region was reduced through the weighting process as these variables were used in weighting adjustments. Sex was also analyzed but did not show significant bias.
    - Analysis 2 – Comparisons of estimates to external totals: Significant differences were found between PIAAC estimates (using final weights) and Labor Force Survey (LFS) quarter 4 of 2011 estimates of employment status. Per Austria, the difference could be caused by the different time spans of the two surveys. In addition, the definition of employment status

differs between LFS and PIAAC, as the latter follows the ILO concept that says "all members of the armed forces, including conscripts, should be defined as being in PAID work." In the LFS, this group of people (armed forces including conscripts) is excluded from the employment analysis.

- Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was above average at 0.43 (0.44 for numeracy). The correlation between the raking dimensions and literacy scores was above average at 0.55 (0.55 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.56 (0.57 for numeracy), which was above the average across countries. Although Austria's response rate was low (53%), this analysis shows that weighting adjustments were effective in reducing NRB because of the high correlation between the survey outcomes and the weighting variables. However, data users need to be cautioned that the analysis is based on correlations between the responding sample (53% of the selected sample) and the weighting variables. That is, the analysis assumes that the same correlations exist for the remaining sampled cases that have no scores (47% of sampled cases).
- Analysis 4 – Comparisons of estimates from alternative weightings: To calculate new weights, the final weighted data were re-calibrated by adding an additional raking dimension. Very small differences were found in the proficiency estimates before and after re-weighting.
- Analysis 5 – Analysis of variables collected during data collection: Austria looked at characteristics of the literacy-related nonrespondents and found that they belonged to the expected sociodemographic groups, except that the low amount of literacy-related cases in one province was unexpected. Bias was reduced by the LRNR weighting adjustment.
- Analysis 6 – Level-of-effort analysis: Austria compared mean literacy scores, as well as age, sex, education, region, urbanization, and nationality, between low level-of-effort cases (interviews conducted with three or fewer contacts) and high level-of-effort cases (interviews conducted with more than three contacts). No significant differences of mean literacy score were found between high level-of-effort and low level-of-effort cases except for the 16-25 years old.
- Analysis 7 – Range of bias: The literacy scores' first plausible value was used to compute the range of scores within the responding sample and predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 64 and the maximum score was 447, for a range of 383. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10^{th} percentile within each weighting cell, and at the other extreme they would all score at the 90^{th} percentile within each weighting cell, the predicted maximum range of the mean was computed to be 48, indicating a minimal potential for bias in outcome statistic. This is a reflection of an effective nonresponse adjustment strategy carried out during weighting. That is,

even though Austria's response rate was low (53%), the effective nonresponse adjustment weighting reduced the potential bias in the outcome statistics to a low level. However, data users need to be cautioned that the analysis is based on assumptions about the range of proficiency scores for sampled cases that have no scores (47% of the sample).

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Austria partially met a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. Austria reached the 7% threshold for 94% of its interviewers.

Austria also partially met a reduced requirement on management. Guidelines 8.1.1B and 8.1.2A required weekly meetings between interviewers and supervisors and an interviewer-supervisor ratio of 20 or less. For the purpose of data evaluation, countries were considered to have met the standard if the meetings between interviewers and supervisors were held every other week and the interviewer-supervisor ratio was 30 or less. Meetings between supervisors and interviewers only occurred on an as-needed basis.

Austria met a reduced requirement on training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. All of Austria's interviewers were provided with at least 15 hours of training. About one-third of interviewers were provided with about 30 hours.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Austria followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

Scoring

To the best of the Consortium's knowledge, Austria followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
  - Core items: 96.0%
  - Literacy Items: 97.9%
  - Numeracy Items: 95.8%
- Scoring reliability of paper-based national booklets
  - Core items: 99.1%
  - Literacy Items: 98.2%
  - Numeracy Items: 98.4%

Assessment data

Overall, 99.0% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Austria, 73.4% of the respondents who completed the BQ took the computer-based cognitive assessment, while 24.2% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Austria, 12.4% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 4.4% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Austria followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Austria. If a respondent started the interview, the likelihood that she/he provided data is at a level above 98% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Austria, about 81.6% of respondents reported income in exact amounts (88.6% across countries) and about 10.9% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Austria, we observed 1.8% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Austria, these percentages were 9.9% for Literacy and 6.8% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Austria, the percentage of nonresponse for Literacy was 5.6%, for Numeracy it was 3.4%, and for PSTRE it was 0.1%.

Canada

Sampling

Canada followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. All QC materials were completed fully.

- Sampling plan: No issues
- Sample selection
  - Home office: The sample selection forms SS-2_DU and SS-2_Person were not submitted until after the data collection period.
  - In field: Canada projected a lack of aboriginal respondents in Yukon and replaced the (not worked) Yukon general sample by an oversample of aboriginals.
- Sample weighting: Canada followed closely to the standards and communicated closely with the Consortium. In order to produce variances that are comparable with other countries and accurately reflect the degrees of freedom for subnational variance estimates using the JK1 approach, Canada implemented a replication approach recommended by the Consortium that is different from the method used in 2003 IALSS. In addition, Canada applied an ad hoc adjustment to integrate the weights/combine all sampled parts (general sample and supplementary samples covering members of the official language minority, individuals between ages 16 and 24, recent immigrants, Aboriginals and Métis).
- Sampling error: Canada’s design effect due to unequal weights is 2.76 for a sample size of 27,285. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 7,848. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (3.45). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. Variation in the weights resulted from some very small initial probabilities of selection and a large number of persons in some households. Further variation was added through nonresponse adjustments. Canada’s targeted number of completed cases was 5 000 in English and 4 500 in French. Respondents could choose to answer PIAAC in either English or French.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 1.8% (residents of Indian reserves, smaller communities in the northern territories, remote and very low population density areas in provinces, non-institutional collective dwellings, other than students in residences).
  - Data collection: Not applicable
- Weighted response rate: 59%
- Nonresponse bias analysis
  - Basic: Canada performed all required analyses using both the general and supplementary samples. Dwelling units located in areas with a higher percentage of individuals having the minority language as a mother tongue showed a lower

response rate. The response rate at the BQ level was higher for women than for men. The non-respondents also tended to live alone or with another individual of the same gender, in apartments, and/or belong to a younger age group (less than 34). All the variables examined in the analyses were used in weighting adjustments.

- Extended: Canada performed all required analyses using both the general and supplementary samples. Their extended analysis provides evidence that bias was reduced through the weighting adjustments.
  - Analysis 1 – Comparisons of estimates before and after weighting: At the screener level, bias in type of dwelling, gender, variables related to household composition, and legal marital status was reduced through weighting. At the BQ level, bias in variables related to household composition, presence of adults having French as a mother tongue in the household, and gender was reduced through weighting.
  - Analysis 2 – Comparisons of estimates to external totals: Some PIAAC estimates (computed using final weights) were outside the confidence intervals produced using the Labor Force Survey for April 2012 data, but there was an overlap between the confidence intervals produced by the two surveys for all industry classification categories.
  - Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was below average at 0.22 (0.23 for numeracy). The correlation between the raking dimensions and literacy scores was above average at 0.53 (0.52 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.54 (0.53 for numeracy), which was about the average across countries. Although the response rate was 59%, this analysis shows an effective reduction in potential NRB due to the moderate correlation between the survey outcomes and the weighting variables. However, data users need to be cautioned that the analysis is based on correlations between the responding sample (59% of the selected sample) and the weighting variables. That is, the analysis assumes that the same correlations exist for the remaining sampled cases that have no scores (41% of sampled cases).
  - Analysis 4 – Comparisons of estimates from alternative weightings: Canada recalibrated to a more basic set of auxiliary variables (province, age, gender, language, immigrant status, and highest level of education). Results calculated using final weights were generally slightly lower than re-weighted proficiency estimates and standard errors were generally smaller, suggesting that additional calibration variables were useful in reducing a potential upward bias in the estimates.
  - Analysis 5 – Analysis of variables collected during data collection: Literacy-related nonrespondents had a specific profile compared to other nonrespondents. They tended to live in apartments, in areas with a lower percentage of individuals being married or living in a common law relationship, in areas with lower median income, in households with more than two adult members, and in households where all adults had a mother

language other than English or French. They tended to be older (aged 55+) and the percentage of women was also higher. Bias was reduced by the LRNR weighting adjustment.

- Analysis 6 – Level-of-effort analysis: Canada defined level-of-effort as the number of days between the first attempt to contact a case and the day of the PIAAC interview. Immigration status and highest level of education completed were characteristics separating low level-of-effort respondents from high level-of-effort respondents. High level-of-effort respondents tended to achieve significantly lower scores than low level-of-effort respondents. There were no significant differences in the distribution of respondents' gender or age.
- Analysis 7 – Range of bias: The literacy scores' first plausible value was used to compute the range of scores within the responding sample and predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 36 and the maximum score was 423, for a range of 386. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10^{th} percentile within each weighting cell, and at the other extreme they would all score at the 90^{th} percentile within each weighting cell, the predicted maximum range of the mean was computed to be 47, indicating a minimal potential for bias in outcome statistics. This is a reflection of an effective nonresponse adjustment strategy carried out during weighting. That is, even though Canada's response rate was low (59%), the effective nonresponse adjustment weighting reduced the potential bias in the outcome statistics to a low level. However, data users need to be cautioned that the analysis is based on assumptions about the range of proficiency scores for sampled cases that have no scores (41% of the sample).

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Canada appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 9.4.2 on interviewer training and Guidelines 8.1.1B and 8.1.2A on management of field staff.

Canada partially met a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of their interviewers, selected randomly, across all dispositions. Canada reached the 7% threshold for 85% of its interviewers. Fifteen percent of interviewers were validated at less than the 7% level.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Canada followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Canada followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 98.3%
    - Literacy Items: 98.3%
    - Numeracy Items: 96.4%
- Scoring reliability of paper-based national booklets
    - Core items: 99.4%
    - Literacy Items: 96.9%
    - Numeracy Items: 98.3%

Assessment data

Overall, 96.6% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Canada, 83.5% of the respondents who completed the BQ took the computer-based cognitive assessment, while 14.7% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Canada, 6.3% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 5.2% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding
To the best of the Consortium's knowledge, Canada followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data
Background data were of very high quality for Canada. If a respondent started the interview, the likelihood that she/he provided data is at a level above 99% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Canada, about 93.4% of respondents reported income in exact amounts (88.6% across countries) and about 2.3% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Canada, we observed 0.9% of cases with breakoffs.

Item nonresponse
Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Canada, these percentages were 13.0% for Literacy and 9.6% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Canada, the percentage of nonresponse for Literacy was 8.6%, for Numeracy it was 6.4%, and for PSTRE it was 0.1%.

Cyprus[13]

Sampling

Cyprus followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. All QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
    - Home office: No issues
    - In field: No field issues detected
- Sample weighting: The Consortium followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights for Cyprus.
- Sampling error: Cyprus' design effect due to unequal weights is 1.39 for a sample size of 5,053. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 2,855. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (1.54). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. Cyprus' sample design involved an equal probability selection at the household level; however, there was variation in the selection probabilities at the person level. Further variation in the weights was added through within-household sampling, nonresponse and calibration adjustments, although the Consortium followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population coverage
    - Frame: The estimated percentage of the target population excluded from the frame was less than 2% (people living in houses built after December 2010)
    - Data collection: Not applicable
- Weighted response rate: 73%
- Nonresponse bias analysis
    - Basic: Cyprus performed all required analyses. Its analysis showed that lower response rates were identified in urban areas and larger districts at the screener level. Potential bias in variables examined were observed also at the screener level for District and Locale, while at the BQ level, statistically significant differences were observed between the respondents and nonrespondents only within Districts. District and Locale have been used in the weighting process (nonresponse adjustment and raking).
    - Extended: The extended analysis provides evidence that bias was reduced through the weighting adjustments. Since Cyprus has a high BQ response rate, analyses 1, 4, and 7 were not required.

[13] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

- Analysis 1 – Comparisons of estimates before and after weighting: Cyprus was not required to do this analysis.
- Analysis 2 – Comparisons of estimates to external totals: PIAAC estimates were compared to Census 2011 and Labor Force Survey 2011 by age, gender, region, education, and employment status. PIAAC estimates are different from Census in the age group 16-19 and the Paphos Urban area.
- Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was below average at 0.21 (0.28 for numeracy). The correlation between the raking dimensions and literacy scores was also below average at 0.39 (0.47 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.39 (0.47 for numeracy), which was below the average across countries. This indicates some potential for reducing NRB due to the correlation between the survey outcomes and the weighting variables.
- Analysis 4 – Comparisons of estimates from alternative weighting: Cyprus was not required to do this analysis.
- Analysis 5 – Analysis of variables collected during data collection: Even though significant differences were found in the distribution by region between the literacy-related cases and the comparison group, these differences cannot be attributed to a possible impact on bias, since for some categories the literacy-related cases are very few. Bias was reduced by the LRNR weighting adjustment.
- Analysis 6 – Level-of-effort analysis: Cyprus compared mean proficiency scores, as well as age, sex, region, urbanization, education, and employment status, between low level-of-effort cases (interviews conducted with five or fewer contacts) and high level-of-effort cases (interviews conducted with more than five contacts). For literacy, low level-of-effort cases were found to have significantly higher proficiency scores than high level-of-effort cases for Nicosia and Larnaca, age group 16-24, and adults with less than upper secondary education. For numeracy, low level-of-effort cases were found to have significantly higher proficiency scores than high level-of-effort cases for Paphos, age group 16-24, adults with less than upper secondary education, and adults out of the labor force.
- Analysis 7 – Range of bias: Cyprus was not required to do this analysis.

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Cyprus generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 10.9.3 on fieldwork validation and Guidelines 8.1.1B and 8.1.2A on management of field staff.

However, Cyprus met a reduced requirement on interviewer training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. Cyprus interviewers were provided with 18 hours of in-person training.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Cyprus followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Cyprus followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 98.3%
    - Literacy Items: 98.8%
    - Numeracy Items: 96.9%
- Scoring reliability of paper-based national booklets
    - Core items: 99.5%
    - Literacy Items: 99.2%
    - Numeracy Items: 98.2%

Assessment data

Overall, 99.9% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Cyprus, 43.7% of the respondents who completed the BQ took the computer-based cognitive assessment, while 38.2% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Cyprus, 28.2% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 2.8% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8%

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Cyprus followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Cyprus. If a respondent started the interview, the likelihood that she/he provided data is at a level above 82% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Cyprus, about 85.7% of respondents reported income in exact amounts (88.6% across countries) and about 3.4% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Cyprus, we observed 17.7% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Cyprus, these percentages were 10.1% for Literacy and 7.5% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Cyprus, the percentage of nonresponse for Literacy was 10.1% and for Numeracy it was 7.1%. Cyprus did not administer the assessment for PSTRE.

The Czech Republic

Sampling

The Czech Republic collected data for two samples: main and supplemental. The target age for the supplemental sample was 16 to 29 year olds, whereas the main sample targeted 16 to 65 year olds. Most QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
    - Home office: The main, reserve, and supplemental samples were selected in a sequential manner, and the selection probabilities provided by the Czech Republic for the reserve and supplemental samples reflected conditional probabilities given the household was not selected for the previous sample. Since the QC sample selection forms were not submitted until after data collection, this was not discovered in time to revise its selection method. Therefore, to create weights that could be used for the combined sample, the Consortium needed to adjust the Czech Republic's probabilities of selection.
    - In field: The Czech Republic used year of birth for screening rather than age or date of birth. This resulted in more cases outside of the target age range. Again, since the QC sample selection forms were not submitted until after data collection, this was not discovered in time to revise its selection method.
- Sample weighting: Selecting the sample in stages required the Consortium to weight the two samples separately and composite them in a final weighting step. Also, using year of birth for screening resulted in 87 persons of age 30 in the supplemental sample. The Czech Republic wanted such cases treated as eligible, so they were weighted with the 29-year-olds.
- Sampling error: The Czech Republic's design effect due to unequal weights is 2.88 for a sample size of 6,102. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 1,725. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (3.53). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. The oversampling of 16 to 29 year olds resulted in variation in the selection probabilities. Further variation in the weights was added through within-household sampling, nonresponse and calibration adjustments, although the Consortium followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population coverage
    - Frame: The estimated percentage of the target population excluded from the frame was 1.8%.
    - Data collection: Not applicable
- Overall weighted response rate: 66%
- Nonresponse bias analysis

- Basic: The Czech Republic performed all required analyses. The decision tree included the following variables as significant predictors of response status at the screener: NUTS (Region); area-level unemployment, gender and age (main sample only); type of municipality; area-level percentage of foreigners (main sample only), household PC and internet connection and educational attainment. Significant predictors at the BQ level according to decision tree analysis are: NUTS (supplemental sample only); municipality; area-level educational attainment, gender, age (main sample only); area-level employment status and entrepreneurs (main sample only).

1) Variables used in the screener level weighting adjustment *for both the main and supplemental samples* included: NUTS (Region); type of municipality; area-level gender, age, unemployment, entrepreneurs and educational attainment (high school); area-level percentage of foreigners and HH PC and internet connection availability.

   Variables used in the BQ level weighting adjustment *for both the main and supplemental samples* included: type of municipality; NUTS (Region); gender; age; area-level unemployment, entrepreneurs, educational attainment (high school), educational attainment (college degree) and HH PC and internet connection availability.

- Extended (preliminary): The Czech Republic performed some, but not all required analyses. Its extended analysis provides evidence that bias was reduced through the weighting adjustments.
  - Analysis 1 – Comparisons of estimates before and after weighting: Potential nonresponse bias in region, age, and gender were reduced through the nonresponse adjustments. However, there were large differences (relative difference > 2) in region and age distributions when comparing the main sample before calibration to the combined sample after calibration.
  - Analysis 2 - Comparisons of estimates to external totals: The Czech Republic compared PIAAC estimates of employment status, reading of books and newspapers, and highest education of father and mother to estimates from the Adult Education Survey (AES) 2011. They also compared PIAAC estimates of household size to European Union – Statistics on Income and Living Conditions (EU-SILC) 2011. Per the Czech Republic, "AES data are significantly different from PIAAC data only in questions such as reading books and newspapers, where even wording and context can influence responses (PIAAC did not stress electronic media and last 12 months)." There were also significant differences in level of education of father and mother and in household size.
  - Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was around the average at 0.35 (0.33 for numeracy). The correlation between the raking dimensions and literacy scores was above average at 0.52 (0.57 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was

0.56 (0.60 for numeracy), which was above the average across countries. This analysis shows an effective reduction in potential NRB due to the high correlation between the survey outcomes and the weighting variables.

- Analysis 4 – Comparisons of estimates from alternative weightings: This analysis was not performed. Per the Czech Republic, “our possibilities to gain another survey data for alternative weighting are rather limited.”
- Analysis 5 – Analysis of variables collected during data collection: An evaluation of the characteristics of literacy-related nonrespondents was not performed because there were a limited number of literacy-related nonrespondents in the Czech Republic. Bias was reduced by the LRNR weighting adjustment.
- Analysis 6 – Level-of-effort analysis: The average literacy score was found to increase with additional visits. The Czech Republic also identified an increase in the percentage employed and differences in age and municipality. This indicates that the thorough data collection efforts helped reduce the bias due to nonresponse.
- Analysis 7 – Range of bias: The literacy scores’ first plausible value was used to compute the range of scores within the responding sample and predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 83 and the maximum score was 445, for a range of 362. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 52, indicating a low potential for bias in outcome statistics. This is a reflection of the higher-than-average response rate (66%) in Czech Republic, combined with an effective nonresponse adjustment carried out during weighting.

Data collection

Based on information provided on QC forms and during monthly QC conference calls, the Czech Republic generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 10.9.3 on fieldwork validation.

The Czech Republic also partially met a reduced requirement on training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. About 75% of the Czech Republic’s interviewers were provided with more than 15 hours; however, about 25% were provided with significantly fewer hours. The Czech Republic offered significantly fewer training hours than recommended on all key aspects (gaining cooperation, BQ administration and assessment administration).

The Czech Republic also partially met a reduced requirement on management. Guidelines 8.1.1B and 8.1.2A required weekly meetings between interviewers and supervisors and an interviewer-supervisor ratio of 20 or less. For the purpose of data evaluation, countries were considered to have met the standard if the meetings between interviewers and supervisors were held every other week and the interviewer-supervisor ratio was 30 or less. Interviewer-supervisor meetings occurred only on an as-needed basis.

Instrument data quality

Translation

To the best of the Consortium's knowledge, the Czech Republic followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, the Czech Republic followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 98.3%
    - Literacy Items: 97.2%
    - Numeracy Items: 96.5%
- Scoring reliability of paper-based national booklets
    - Core items: 100.0%
    - Literacy Items: 99.6%
    - Numeracy Items: 100.0%

Assessment data

Overall, 99.8% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In the Czech Republic, 74.4% of the respondents who completed the BQ took the computer-based cognitive assessment, while 24.4% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In the Czech Republic, 13.5% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 2.4% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries,

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding
To the best of the Consortium's knowledge, the Czech Republic followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data
Background data were of very high quality for the Czech Republic. If a respondent started the interview, the likelihood that she/he provided data is at a level above 99% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In the Czech Republic, about 83.4% of respondents reported income in exact amounts (88.6% across countries) and about 5.2% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In the Czech Republic, we observed 0.6% of cases with breakoffs.

Item nonresponse
Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In the Czech Republic, these percentages were 5.9% for Literacy and 3.7% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in the Czech Republic, the percentage of nonresponse for Literacy was 5.9%, for Numeracy it was 3.3%, and for PSTRE it was 0.0%.

Denmark

Sampling

Denmark followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. Most QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
  - Home office: The sample selection form was not submitted prior to the data collection period. One source of attrition is due to 14.5 of the Danish population aged 16-65 years who are registered in a so-called opt-out register. That is, they have informed the authorities that their names, addresses, and phone numbers must not be given to research institutions, etc., wanting to contact them for an interview. Only persons without researcher protection can be contacted. Statistics Denmark was able to get all required register information regarding the persons in the opt-out register, and included them for the weighting and nonresponse bias analysis.
  - In field: Not applicable
- Sample weighting: Denmark followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights. Not all could be verified, including:
  - The comparison with alternative external totals was not done by Denmark, therefore, we were not able to validate the totals. However, Denmark has registered information and adjusted the weights to reflect the population totals. Therefore, they found it superfluous to check the totals against alternative external totals—the source would be the same in most cases or the quality would be much lower.
- Sampling error: Denmark's design effect due to unequal weights is 1.27 for a sample size of 7,328. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 5,861. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (1.24). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. The sample design involved an oversample of immigrants and adults 55-65 years old. Further variation in the weights was added through nonresponse and calibration adjustments, although they followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population Coverage
  - Frame: The estimated percentage of the target population excluded from the frame was less than 0.1% (undocumented immigrants).
  - Data collection: The weighted percentage of cases excluded because they were inaccessible was 5.0%.
- Weighted response rate: 50%
- Nonresponse bias analysis

- Basic: Denmark performed all required analyses. Prior to weighting adjustments, age, region, education, and employment show significant potential for bias. In particular, overrepresentation occurred for older adults, regions close to the capital, people with higher education, employed people, and students. All required variables were used in the NRBA, as well as income, type of family, ethnicity, and mobility. The logistic regression was done by strata (age group). Within the 10 strata, there are several indications of the potential for bias. Region, education level, and mobility showed significant effects for at least five of the strata. Logistic regressions show that non-weighting variables of disposable income and average family income had a significant potential for bias for a small number of age groups. However, these are likely correlated with gross income, which was used in weighting.
- Extended: Denmark performed all required analyses. Its extended analysis provides evidence that bias was reduced through the weighting adjustments.
    - Analysis 1 – Comparisons of estimates before and after weighting: There were very small standard errors, which may lead to more statistically significant results. The calibrated weights reflect the population distribution according to the registers. The base weights on the eligible sample gives the same picture, indicating that the eligible sample represents the population. The nonresponse pattern results in skewed estimates and thus substantial possibility for nonresponse bias. However, the nonresponse adjusted weight, to a large extent, remedies this. For variables not used in the weighting, the base weights for the eligible sample gave the same picture as the calibrated weights. The nonresponse pattern results in different estimates, and thus substantial possibility for nonresponse bias. However, the nonresponse adjusted weight, to a large extent, remedies this.
    - Analysis 2 – Comparisons of estimates to external totals: The external totals table shows differences between the PIAAC estimates on age and income using the final calibrated weight, with estimates from the registry. Significant differences were found for age groups 16-20 (higher in PIAAC), 21-25 (lower), 56-60 (lower), 61-65 (higher); and in low income (lower). Since the final weights were calibrated by age group using registry totals, it is a bit surprising, although it was done for different categories of age than what was used for calibration.
    - Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse-related variables and literacy scores was above average at 0.47 (0.42 for numeracy). The correlation between the raking dimensions and literacy scores was slightly below average at 0.43 (0.39 for numeracy). The correlation between literacy scores and the combination of nonresponse related variables and raking dimensions was 0.50 (0.46 for numeracy), which was about the average across countries. This indicates some potential for reducing NRB due to the moderate correlation between the survey outcomes and the weighting variables. The analysis shows that weighting adjustments were moderately effective in reducing NRB because of the correlation between the survey

outcomes and the weighting variables. However, data users need to be cautioned that the analysis is based on correlations between the responding sample (50% of the selected sample) and the weighting variables. That is, the analysis assumes that same correlations exist for the remaining sampled cases that have no scores (50% of sampled cases).

- Analysis 4 – Comparisons of estimates from alternative weightings: The standard errors on the estimates are small and of the same size in both cases.
    - The estimates themselves are larger in the re-weighting. Denmark expected nonresponse bias to result in overestimation of the proficiency scores, and the re-weighting results support the hypothesis that the more elaborated calibration model used in PIAAC weighting reduces bias the most.
- Analysis 5 – Analysis of variables collected during data collection: Region, gender, and age groups all showed differences, however, the Denmark weighting procedures separated the LRNR cases, therefore treating them appropriately.
- Analysis 6 – Level-of-effort analysis: No differences were found between men and women in the level-of-effort needed to attain response. Differences between the regions were found in the level-of-effort needed to attain response. In the Sealand region more than half of the responses were attained with low level-of-effort, whereas in the other regions it was around 40%. Differences between the age groups were found in the level-of-effort needed to attain response. The overall trend being “the younger the higher level-of-effort needed.” The most difficult group to attain response from was however the 25-34 year-olds. There was a tendency toward a higher PVLIT1-score among low level-of-effort part of the citizens in the capital. Also, a tendency toward a higher PVLIT1-score among the high level-of-effort part of the 35-44 year-olds was seen. In general, such differences between low and high level-of-effort indicates some reduction in nonresponse bias.
- Analysis 7 – Range of bias: The literacy scores’ first plausible value was used to compute the range of scores within the responding sample and predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 38 and the maximum score was 405, for a range of 366. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 50, indicating a low potential for bias in outcome statistics. This is a reflection of an effective nonresponse adjustment strategy carried out during weighting. That is, even though Denmark’s response rate was low (50%), the effective nonresponse adjustment weighting reduced the potential bias in the outcome statistics to a low level. However, data users need to be cautioned that the analysis

is based on assumptions about the range of proficiency scores for sampled cases that have no scores (50% of the sample).

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Denmark generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 10.9.3 on fieldwork validation.

Denmark met a reduced requirement on interviewer training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. About half of Denmark's interviewers were provided with a minimum of 15 hours of training.

Denmark met a reduced requirement on management. Guidelines 8.1.1B and 8.1.2A required weekly meetings between interviewers and supervisors and an interviewer-supervisor ratio of 20 or less. For the purpose of data evaluation, countries were considered to have met the standard if the meetings between interviewers and supervisors were held every other week and the interviewer-supervisor ratio was 30 or less. In Denmark, supervisor assignments were between 20 and 30 interviewers.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Denmark followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Denmark followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 97.1%
    - Literacy Items: 97.3%
    - Numeracy Items: 95.9%
- Scoring reliability of paper-based national booklets

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

- Core items: 99.7%
- Literacy Items: 98.9%
- Numeracy Items: 99.3%

Assessment data

Overall, 97.4% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Denmark, 87.1% of the respondents who completed the BQ took the computer-based cognitive assessment, while 11.8% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Denmark, 5.7% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 4.3% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium’s knowledge, Denmark followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Denmark. If a respondent started the interview, the likelihood that she/he provided data is at a level above 99% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Denmark, about 96.3% of respondents reported income in exact amounts (88.6% across countries) and about 1.0% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Denmark, we observed 0.4% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Denmark, these percentages were 18.0% for Literacy and 10.3% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Denmark, the percentage of nonresponse for Literacy was 7.9%, for Numeracy it was 5.4%, and for PSTRE it was 0.3%.

England/Northern Ireland (UK)

Sampling

England/Northern Ireland (UK) followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting.

- Sampling plan: No issues
- Sample selection
    - Home office: No issues
    - In field: The theoretical person base weights (THEOR_PBWT) were derived from imputed values of the number of eligible people in the sampled household (NUM_ELG) for 52 cases (49 in England (UK) and 3 in Northern Ireland (UK)) due to a technical problem with the contact data that the interviewers entered.
- Sample weighting: The Consortium followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create two sets of weights separately for England (UK) and Northern Ireland (UK). England/Northern Ireland (UK) did not collect age and gender for all sampled persons during the screener. A special adjustment was implemented so that literacy-related nonrespondents with age and gender successfully collected represented those with age or gender not successfully collected.
- Sampling error: The design effect due to unequal weights is 1.35 for England (UK) for a sample size of 5,131; and 1.54 for Northern Ireland (UK) for a sample size of 3,761. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 2 176 for England (UK) and 563 for Northern Ireland (UK). The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (2.33 for England (UK) and 6.62 for Northern Ireland (UK)). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. England/Northern Ireland (UK)'s address sample was an equal probability sample in both England (UK) and Northern Ireland (UK). Variation in the selection probabilities was introduced from (a) subsampling households for addresses containing multiple households, and (b) the within-household selection at the person level. Further variation in the weights was added through nonresponse and calibration adjustments, although the Consortium followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population Coverage
  - Frame: The combined estimated percentage of the target population excluded from the frame in England (UK) and Northern Ireland (UK) was 2.0% (individuals living in private residences that are not listed on the ”residential” version of the Postal Address File; or, in Northern Ireland (UK), not listed on the NI(POINTER) database).
  - Data collection: Not applicable
- Weighted response rate: 59% for England (UK); 65% for Northern Ireland (UK)
- Nonresponse bias analysis
  - Basic: England/Northern Ireland (UK) performed all required analyses and used all required analysis variables (neighborhood characteristics).
    - England (UK): The screener response rate varied by region (from 77% in London to 89% in North East England (UK)). The highest category screener response rate was 90% (Output Area Classification: terraced blue collar neighborhoods) and the lowest was 74% (Output Area Classification: transient communities). The highest category BQ response rate was 84% (third quintile category of % Indian) and the lowest was 59% (Output Area Classification: transient communities). The lowest regional BQ response rate was in London (61%). Screened households differed from nonscreened households in terms of neighborhood profile. Neighborhoods with a high proportion of residents not born in the UK or of Black or Bangladeshi descent were underrepresented in the screened household sample. London was also underrepresented. Neighborhoods with a high proportion of Black residents were underrepresented in the BQ respondent sample while neighborhoods with a high proportion of older people (aged 65+) and of those with a caring responsibility were overrepresented. London was underrepresented. The classification tree found that region was the only significant screener response rate predictor. The classification tree identified the proportion aged 65+ as the only significant BQ response rate predictor. BQ response rates tended to be higher in neighborhoods with an older-than-average age profile.
    - Northern Ireland (UK): The highest category screener response rate was 90% (Output Area Classification: senior communities) and the lowest was 25% (Output Area Classification: Asian communities). Excluding this very small sample size category, the lowest was 70% (Output Area Classification: public housing). It is notable that the screener response rate in the capital Belfast was only 72%. The highest category BQ response rate - excluding categories with small sample sizes - was 86% (Output Area Classification: young families in terraced homes and the lowest was 65% (Output Area Classification: village life). The lowest regional BQ response rate was in the North (75%). Neighborhoods in Belfast were the most underrepresented in the screened household sample. There were no significant profile differences between BQ responders and nonresponders in Northern Ireland (UK). The classification tree identified region and the proportion aged 65+ as strong discriminators of screener response rates.

The classification tree identified the proportion aged 65+ as the strongest predictor of BQ response rates.

- Extended: England/Northern Ireland (UK) did not perform all required analyses. Although some paradata were collected, the three agencies responsible for fieldwork did not collect them in a consistent fashion so that it could be used for analytical purpose.
  - Analysis 1 – Comparisons of estimates before and after weighting: In England (UK) and Northern Ireland (UK), at both the screener and BQ levels, bias in region was reduced through the weighting process as it was used in weighting adjustments. The base-weighted respondent profile was very similar to the base-weighted sampled person profile.
  - Analysis 2 – Comparisons of estimates to external totals: In both England (UK) and Northern Ireland (UK), large differences were found between PIAAC estimates (using final weights) and Census 2011 totals of employment status, ethnic group and general health. In Northern Ireland (UK), nontrivial differences were also found for qualification. (In England (UK), Census 2011 data on qualification will not be released until late August, 2013.) For age, a large difference was found for category 60-65 (higher in PIAAC), which came as a surprise given the PIAAC control totals were based on census totals updated by birth, death, and immigration/emigration data. England/Northern Ireland (UK) noted that disparities between the PIAAC estimates and Census 2011 totals may reflect the difference in the interview mode (interviewer-assisted vs. self-administered).
  - Analysis 3 – Correlation of auxiliary variables and proficiency estimates:
    - England (UK): The correlation between the BQ nonresponse cells and literacy scores was below average at 0.32 (0.35 for numeracy). The correlation between the raking dimensions and literacy scores was average at 0.48 (0.51 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.52 (0.56 for numeracy), which was about the average across countries. Although the response rate was 59%, this analysis shows an effective reduction in potential NRB due to the moderate correlation between the survey outcomes and the weighting variables. However, data users need to be cautioned that the analysis is based on correlations between the responding sample (59% of the selected sample) and the weighting variables. That is, the analysis assumes that the same correlations exist for the remaining sampled cases that have no scores (41% of sampled cases).
    - Northern Ireland (UK): The correlation between the BQ nonresponse cells and literacy scores was below average at 0.33 (0.36 for numeracy). The correlation between the raking dimensions and literacy scores was above average at 0.55 (0.58 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.57 (0.60

for numeracy), which was higher than average across countries. Although the response rate was 65%, this analysis shows an effective reduction in potential NRB due to the high correlation between the survey outcomes and the weighting variables. However, data users need to be cautioned that the analysis is based on correlations between the responding sample (65% of the selected sample) and the weighting variables. That is, the analysis assumes that the same correlations exist for the remaining sampled cases that have no scores (35% of sampled cases).

- Analysis 4 – Comparisons of estimates from alternative weightings: This analysis was not performed.
- Analysis 5 – Analysis of variables collected during data collection: This analysis was not performed due to limited data on nonrespondents: gender in most cases (when the household was screened) but very rarely age for nonrespondents.
- Analysis 6 – Level-of-effort analysis: This analysis was not performed due to the lack of consistent paradata on the number of visits per case. Date of interview could not be used because the sample was released in batches (and it was more of a drip-feed approach in Northern Ireland (UK)). The alternative approach England/Northern Ireland (UK) took was to report the correlation between the effective "response factor" and the proficiency scores, showing that the lower the response propensity, the lower the proficiency score. This suggests a slight upwards bias may remain in the estimates, reflecting the partial, not total ability of calibration to counter nonresponse bias.
- Analysis 7 – Range of bias:
  - England (UK): The response rate for England (UK) was 59%. The Literacy scores' first plausible value was used to compute the range of scores within the responding sample and to predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 84 and the maximum score was 409, for a range of 325. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 47, indicating a minimal potential for bias in outcome statistics. This is a reflection of an effective nonresponse adjustment strategy carried out during weighting. That is, even though England (UK)'s response rate was low (59%), the effective nonresponse adjustment weighting reduced the potential bias in the outcome statistics to a low level. However, data users need to be cautioned that the analysis is based on assumptions about the range of proficiency scores for sampled cases that have no scores (41% of the sample).
  - Northern Ireland (UK): The response rate for Northern Ireland (UK) was 65%. The Literacy scores' first plausible value was used

to compute the range of scores within the responding sample and to predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 97 and the maximum score was 419, for a range of 322. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 37, indicating a minimal potential for bias in outcome statistics. This is a reflection of the relatively high response rate (65%) in Northern Ireland (UK), combined with an effective nonresponse adjustment carried out during weighting.

Data collection

England (UK)

Based on information provided on QC forms and during monthly QC conference calls, England (UK) generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Guidelines 8.1.1B and 8.1.2A on management of field staff.

England (UK) partially met a reduced requirement on training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. Interviewers were provided with about 10 hours of in-person training and were offered significantly fewer training hours than recommended on key aspects (gaining cooperation and assessment administration). However, interviewers were experienced and had previously received general interviewing techniques training and at-home project-specific training.

England (UK) did not meet a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. England (UK) reached the 7% threshold for 20% of its interviewers. Eighty percent of interviewers were validated at less than the 7% level. However, at least 10% of cases were validated overall.

Northern Ireland (UK)

Based on information provided on QC forms and during monthly QC conference calls, Northern Ireland (UK) generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Guidelines 8.1.1B and 8.1.2A on management of field staff.

Northern Ireland (UK) partially met a reduced requirement on training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the

Consortium. Interviewers were provided with about 10 hours of in-person training and were offered significantly fewer training hours than recommended on key aspects (gaining cooperation and assessment administration). However, interviewers were experienced and had previously received general interviewing techniques training and at-home project-specific training.

Northern Ireland (UK) partially met a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. Northern Ireland (UK) reached the 7% threshold for 95% of its interviewers. Five percent of interviewers were validated at less than the 7% level.

Instrument data quality

Translation

To the best of the Consortium's knowledge, England/Northern Ireland (UK) followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, England/Northern Ireland (UK) followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring international anchor booklets
    - Core items: 98.4%
    - Literacy Items: 98.8%
    - Numeracy Items: 96.6%
- Scoring reliability of paper-based national booklets
    - Core items: 100.0%
    - Literacy Items: 100.0%
    - Numeracy Items: 100.0%

Assessment data

Overall, 97.4% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In England/Northern Ireland (UK), 83.4% of the

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

respondents who completed the BQ took the computer-based cognitive assessment, while 14.1% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In England/Northern Ireland (UK), 4.8% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 5.8% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, England/Northern Ireland (UK) followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for England/Northern Ireland (UK). If a respondent started the interview, the likelihood that she/he provided data is at a level above 98% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In England/Northern Ireland (UK), about 89.8% of respondents reported income in exact amounts (88.6% across countries) and about 2.8% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In England/Northern Ireland (UK), we observed 1.4% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In England/Northern Ireland (UK), these percentages were 10.5% for Literacy and 7.2% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in England/Northern Ireland (UK), the percentage of nonresponse for Literacy was 7.2%, for Numeracy it was 5.5%, and for PSTRE it was 0.1%.

Estonia

Sampling

Estonia followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. All QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
  - Home office: No issues
  - In field: Not applicable
- Sample weighting: The Consortium followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights for Estonia.
- Sampling error: Estonia's design effect due to unequal weights is 1.04 for a sample size of 7,632. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 3 785. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (2.00). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 2.8% (undocumented immigrants and people without a detailed address).
  - Data collection: The weighted percentage of cases excluded because they were inaccessible was 0.6%.
- Weighted response rate: 63%
- Nonresponse bias analysis
  - Basic: Estonia performed all required analyses. Its basic analysis showed significantly low response rates for males, 26-35 year olds, people with non-Estonian mother tongue, several counties, big city, and areas with higher education. Age, gender, mother tongue, counties, urbanization, education, and unemployment were used in weighting adjustment.
  - Extended: The extended analysis provides evidence that bias was reduced through the weighting adjustments.
    - Analysis 1 – Comparisons of estimates before and after weighting: Bias in age, gender, mother tongue, urbanization, county, area-level education and unemployment was reduced through the weighting process as these variables were used in weighting adjustments. No other variables were analyzed.
    - Analysis 2 – Comparisons of estimates to external totals: PIAAC estimates were compared to Census 2011 by age, gender, county, and area-level unemployment. PIAAC estimates are larger than Census both overall and for most of the domains compared. This is probably due to the fact that

PIAAC estimates are based on Population Register, which includes people who moved to other countries, while Census has some undercoverage.

- Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was below average at 0.31 (0.30 for numeracy). The correlation between the raking dimensions and literacy scores was also below average at 0.31 (0.29 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.37 (0.35 for numeracy), which was below the average across countries. That is, weighting adjustments were not as effective in reducing bias, as compared to other countries, because of the lower-than-average correlation between the survey outcomes and the weighting variables. However, Estonia had a higher-than-average response rate (63%), as compared to other countries, implying that the potential for bias is likely to be somewhat lower as compared to countries with lower response rates. This indicates some potential for reducing NRB due to the moderate correlation between the survey outcomes and the weighting variables.
- Analysis 4 – Comparisons of estimates from alternative weightings: To calculate new weights, the final weighted data were recalibrated by the area-level percent of unemployment. Very small differences were found in the proficiency estimates before and after reweighting.
- Analysis 5 – Analysis of variables collected during data collection: Estonia will not perform this analysis since they do not have any additional information besides disposition codes and its proportion of literacy-related cases is very low (0.3%). Bias was reduced by the LRNR weighting adjustment.
- Analysis 6 – Level-of-effort analysis: Estonia compared mean proficiency scores, as well as age, sex, mother tongue, urbanization, county, area-level education and unemployment, between low level-of-effort cases (interviews conducted with five or fewer contacts) and high level-of-effort cases (interviews conducted with more than five contacts). High level-of-effort cases were found to have significantly lower proficiency scores than low level-of-effort cases for females, 16-25 years old, 36-45 years old, and several counties, suggesting a high amount of contact should be carried out.
- Analysis 7 – Range of bias: The literacy scores' first plausible value was used to compute the range of scores within the responding sample and predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 91 and the maximum score was 406, for a range of 315. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 41, indicating a minimal potential for bias in outcome statistics. This is a reflection of the higher-than-average response rate (63%) in Estonia. That is, as a result of achieving a

higher response rate, the potential for remaining bias is minimal even though the weighting adjustments were not as effective, as compared to other countries, in reducing bias in outcome statistics.

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Estonia generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 10.9.3 on fieldwork validation and Guidelines 8.1.1B and 8.1.2A on management of field staff.

Estonia met a reduced requirement on interviewer training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. Estonian interviewers were provided with at least 24 hours of in-person training.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Estonia followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Estonia followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 95.5%
    - Literacy Items: 95.5%
    - Numeracy Items: 95.5%
- Scoring reliability of paper-based national booklets
    - Core items: 99.5%
    - Literacy Items: 97.9%
    - Numeracy Items: 98.7%

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

Assessment data

Overall, 99.0% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Estonia, 70.7% of the respondents who completed the BQ took the computer-based cognitive assessment, while 28.5% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Estonia, 17.4% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 3.5% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Estonia followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Estonia. If a respondent started the interview, the likelihood that she/he provided data is at a level above 99% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Estonia, about 82.0% of respondents reported income in exact amounts (88.6% across countries) and about 1.5% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Estonia, we observed 0.4% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Estonia, these percentages were 9.2% for Literacy and 6.6% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Estonia, the percentage of nonresponse for Literacy was 7.8%, for Numeracy it was 4.8%, and for PSTRE it was 0.1%.

Finland

Sampling

Finland followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. All QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
  - Home office: No issues
  - In field: Not applicable
- Sample weighting: The Consortium followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights for Finland.
- Sampling error: Finland's design effect due to unequal weights is 1.05 for a sample size of 5,464. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 5 464. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (0.94). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. The effective sample size is set equal to the actual number of cases with plausible values since the overall design effect is less than 1.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 0.2% (undocumented immigrants and asylum seekers).
  - Data collection: The weighted percentage of cases excluded because they are inaccessible was 0.5%.
- Weighted response rate: 66%
- Nonresponse bias analysis
  - Basic: Finland performed all required analyses. However, employment was not included in the analysis and they didn't state why. Its analysis showed significantly lower response propensities for people with lower education (0) and urban municipalities, and higher response rates among high education groups (5-6), adults age 56-64, Swedish speakers and rural municipalities. Logistic regression also shows significant influence of region and family status. Age, gender, education, native language, major region and urbanism were used in weighting adjustments.
  - Extended: Finland performed analyses of comparisons of before and after weighting adjustments, comparisons of weighted estimates to external totals, correlation of auxiliary variables and proficiency estimates, literacy-related disposition codes, level-of-effort analysis, and calculation of the range of potential bias. The extended analysis provides evidence that bias was reduced through the weighting adjustments.
    - Analysis 1 – Comparisons of estimates before and after weighting: Relative difference was reduced for all categories. Among those with

relative difference>2 and absolute difference>1 before weighting categories, ages 56-64, education and urbanism's relative difference was reduced to less than two, while language's difference was reduced to less than 0.11 (relative difference was still large due to low variance).

- Analysis 2 – Comparisons of estimates to external totals: Estimates of age and gender are consistent with the registry. The unemployed and not in Labor force counts are significantly different from that in the Labor Force Survey. This could be because of differences in definition and questionnaire structure.
- Analysis 3 – The correlation between the BQ nonresponse cells and literacy scores was 0.53 (0.50 for numeracy). The correlation between the raking dimensions and literacy scores was 0.59 (0.56 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.60 (0.58 for numeracy), which was above the average across countries. The analysis shows that weighting adjustments were effective in reducing NRB because of the high correlation between the survey outcomes and the weighting variables.
- Analysis 4 – Comparisons of estimates from alternative weighting: Per Finland, "Various alternative weights were tried during the data collection phase and finally we ended up to the current ones: no improvement could be achieved with other potential variables."
- Analysis 5 – Analysis of variables collected during data collection: Finland looked into literacy-related nonrespondents and found they are mostly less educated and speak other languages. They are also more likely to live in capital areas and urban municipalities. The result shows that bias was reduced by the LRNR weighting adjustment.
- Analysis 6 – Level-of-effort analysis: Finland defined level-of-effort using three variables: number of contacts, time gap between the first attempt and the last, and a proxy for overall effort (standardized principal component of 1 and 2 above plus information on how many interviewers were assigned to handle the problematic cases). A binary indicator was created for each factor using a cut-off point at the third quartile. There are significant differences in the distribution— late respondents have a higher percentage of young, live in the southern parts and urban areas, or speak a different language. Late respondents tend to have higher score, although the difference is not significant.
- Analysis 7 – Range of bias: The literacy scores' first plausible value was used to compute the range of scores within the responding sample and predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 30 and the maximum score was 441, for a range of 411. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10^{th} percentile within each weighting cell, and at the other extreme they would all score at the 90^{th} percentile within each weighting cell, the predicted maximum range of the mean was computed to be 35, indicating a minimal potential for bias in outcome statistics. This is a reflection of the relatively high

response rate (66%) in Finland, combined with an effective nonresponse adjustment step carried out during weighting.

Data collection

Finland met a reduced requirement on training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. Finland's interviewers were provided with 15 hours of training.

Finland met a reduced requirement on management. Guidelines 8.1.1B and 8.1.2A required weekly meetings between interviewers and supervisors and an interviewer-supervisor ratio of 20 or less. For the purpose of data evaluation, countries were considered to have met the standard if the meetings between interviewers and supervisors were held every other week and the interviewer-supervisor ratio was 30 or less. Meetings between supervisors and interviewers occurred at least every two weeks and the interviewer-supervisor ratio was between 20 and 30.

Finland did not meet a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. Finland reached the 7% threshold for 46% of its interviewers. Fifty-four percent of interviewers were validated at less than the 7% level.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Finland followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Finland followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 97.5%

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

    - Literacy Items: 98.4%
    - Numeracy Items: 96.1%
- Scoring reliability of paper-based national booklets
    - Core items: 99.8%
    - Literacy Items: 96.4%
    - Numeracy Items: 98.9%

Assessment data

Overall, 97.2% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Finland, 83.2% of the respondents who completed the BQ took the computer-based cognitive assessment, while 16.2% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Finland, 10.0% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 3.6% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Finland followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Finland. If a respondent started the interview, the likelihood that she/he provided data is at 100% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Finland, about 93.5% of respondents reported income in exact amounts (88.6% across countries) and about 3.7% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Finland, we observed 0.0% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Finland, these percentages were 11.2% for Literacy and 8.3% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Finland, the percentage of nonresponse for Literacy was 4.7%, for Numeracy it was 3.0%, and for PSTRE it was 0.4%.

Flanders (Belgium)

Sampling

Flanders (Belgium) followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. All QC materials were completed fully.

- Sampling plan: No issues
- Sample selection
  - Home office: No issues
  - In field: Not applicable
- Sample weighting: The Consortium followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights for Flanders (Belgium). An unknown eligibility adjustment was not needed because there were no inaccessible cases with unknown whereabouts. A literacy-related nonresponse adjustment was also not needed because all literacy-related nonrespondents had age and gender collected.
- Sampling error: Flanders (Belgium)'s sample design involved an equal probability sample. The design effect due to unequal weights is 1.04 for a sample size of 5,463. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 3,215. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (1.55). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 1.0% (undocumented immigrants).
  - Data collection: The weighted percentage of cases excluded because they are inaccessible was 4.0%.
- Weighted response rate: 62%
- Nonresponse bias analysis
  - Basic: Flanders (Belgium) performed all required analyses. The required variables education and employment from the Labor Force Survey were not included in all required analyses. Its analysis showed a lower response rate for 26 to 35 year olds and males, based on registry information. Respondents and nonrespondents were significantly different by age, gender, province, employment status and educational attainment. The classification tree analysis indicated that there was a large proportion of nonrespondents in Vlaams Brabant (13.9%) due to literacy related reasons. This result was expected because of the large proportions of French-speaking Flemings and foreign speakers in Vlaams Brabant. Age, gender, and province were used in weighting adjustments.
  - Extended: Flanders (Belgium) performed all required analyses. Its extended analysis provides evidence that bias was reduced through the weighting adjustments.

- Analysis 1 – Comparisons of estimates before and after weighting: Flanders (Belgium) examined age, gender and province. Bias in age and province was reduced through the weighting adjustments as these variables were used in weighting. No bias was found in gender before the weighting adjustments.
- Analysis 2 – Comparisons of estimates to external totals: Large differences were found between the PIAAC estimates (computed using final weights) and the 2011 Labor Force Survey estimates of age and educational attainment. Flanders (Belgium) did not provide an explanation for the differences.
- Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was below average at 0.32 (0.33 for numeracy). The correlation between the raking dimensions and literacy scores was below average at 0.33 (0.33 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.36 (0.36 for numeracy), which was below the average across countries implying that weighting adjustments were not as effective in reducing NRB, as compared to other countries, on average. This is due to the low correlation between the survey outcomes and the weighting variables. However, the Flanders (Belgium) response rate (62%) was slightly higher than the average response rate, implying that the potential for bias is lower as compared to countries with lower response rates.
- Analysis 4 – Comparisons of estimates from alternative weightings: To compute alternative weights, the final weighted data were recalibrated to employment status and educational attainment, which were not available at the time of weighting. Although results calculated using final weights were generally slightly lower than re-weighted proficiency estimates, the estimates were very similar.
- Analysis 5 – Analysis of variables collected during data collection: Flanders (Belgium) compared literacy-related nonrespondents with non-literacy-related nonrespondents on age, gender and province and found a large proportion of literacy-related nonrespondents in Vlaams Brabant. This result was expected and confirmed its finding from the basic analysis. Bias was reduced by the LRNR weighting adjustment.
- Analysis 6 – Level-of-effort analysis: Flanders (Belgium) defined level-of-effort by the number of attempts to contact (no reconversion= less than 5 attempts vs. reconversion= 5 or more attempts). There was significant difference between the two level-of-effort groups when controlling for province. There were no significant differences in the distribution of respondents' age, gender, employment status and educational attainment. High level-of-effort respondents generally achieved lower scores than low level-of-effort respondents. Easier-to-contact men had higher proficiency scores than difficult-to-contact men. Easier-to-contact respondents in Limburg had significantly higher proficiency scores than difficult-to-contact participants from Limburg. Easier-to-contact participants with jobs

had significantly higher proficiency scores than difficult-to-contact participants with jobs. Easier-to-contact respondents with ISCED 3 or ISCED 4 qualifications had significantly higher proficiency scores than difficult-to-contact respondents with the same educational level.

- Analysis 7 – Range of bias: The literacy scores' first plausible value was used to compute the range of scores within the responding sample and predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 89 and the maximum score was 441, for a range of 323. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 47, indicating a minimal potential for bias in outcome statistics. This is a reflection of the higher-than-average response rate (62%) in Flanders (Belgium). That is, as a result of achieving a higher response rate, the potential for remaining bias is minimal even though the weighting adjustments were not as effective, as compared to other countries, in reducing bias in outcome statistics.

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Flanders (Belgium) generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 9.4.2 on interviewer training.

Flanders (Belgium) partially met a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. Flanders (Belgium) reached the 7% threshold for 84% of its interviewers.

Flanders (Belgium) met a reduced requirement on management. Guidelines 8.1.1B and 8.1.2A required weekly meetings between interviewers and supervisors and an interviewer-supervisor ratio of 20 or less. For the purpose of data evaluation, countries were considered to have met the standard if the meetings between interviewers and supervisors were held every other week and the interviewer-supervisor ratio was 30 or less. Flanders (Belgium)'s supervisor assignments included 25 interviewers.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Flanders (Belgium) followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Flanders (Belgium) followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
  - Core items: 99.0%
  - Literacy Items: 97.8%
  - Numeracy Items: 95.8%
- Scoring reliability of paper-based national booklets
  - Core items: 99.7%
  - Literacy Items: 99.4%
  - Numeracy Items: 99.4%

Assessment data

Overall, 99.2% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Flanders (Belgium), 78.7% of the respondents who completed the BQ took the computer-based cognitive assessment, while 15.2% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Flanders (Belgium), 5.3% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 3.7% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Flanders (Belgium) followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Flanders (Belgium). If a respondent started the interview, the likelihood that she/he provided data is at a level above 94% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Flanders (Belgium), about 84.2% of respondents reported income in exact amounts (88.6% across countries) and about 10.6% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Flanders (Belgium), we observed 5.2% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Flanders (Belgium), these percentages were 11.1% for Literacy and 6.0% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Flanders (Belgium), the percentage of nonresponse for Literacy was 6.9%, for Numeracy it was 4.9%, and for PSTRE it was 0.0%.

France

Sampling

France followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. All QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues.
- Sample selection
  - Home office: No issues
  - In field: Not applicable
- Sample weighting: The Consortium followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights for France.
- Sampling error: France's design effect due to unequal weights is 1.05 for a sample size of 6,993. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 6,867. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (1.01). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. France selected the sample in two stages, and in the first stage the Interviewer Action Areas (IAAs) were selected using a balanced sampling design. During the weighting process, Westat used the approximate variance estimator for balanced samples proposed by Deville and Tille (2005) and followed Fay's method (1984) to generate 80 replicate weights.

Coverage and Nonresponse Bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was less than 2.6% (young adults who never claimed any income and are not attached to its parents' households (0.6%) and undocumented immigrants (less than 2%)).
  - Data collection: The weighted percentage of cases excluded because they were inaccessible was 1.4%.
- Weighted response rate: 67%
- Nonresponse bias analysis
  - Basic: France performed all required analyses. The chi-square analysis showed differential response rates by age, region, and income.
  - Extended: France did not complete all the required analyses.
    - Analysis 1 – Comparisons of estimates before and after weighting: Bias in age, gender, region and income was reduced through the weighting process as these variables were used in weighting adjustments.
    - Analysis 2 – Comparisons of estimates to external totals: Was not performed.
    - Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was above average at 0.44 (0.46 for numeracy). The correlation between the

raking dimensions and literacy scores was also above average at 0.57 (0.61 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.60 (0.64 for numeracy), which was above the average across countries. This analysis shows that weighting adjustments were effective in reducing NRB because of the high correlation between the survey outcomes and the weighting variables. However, data users need to be cautioned that the analysis is based on correlations between the responding sample (67% of the selected sample) and the weighting variables. That is, the analysis assumes that same correlations exist for the remaining sampled cases that have no scores (33% of the sampled cases).

- Analysis 4 – Comparisons of estimates from alternative weightings: Was not performed.
- Analysis 5 – Analysis of variables collected during data collection: Was not performed.
- Analysis 6 – Level-of-effort analysis: Was not performed.
- Analysis 7 – Range of bias: The literacy scores' first plausible value was used to compute the range of scores within the responding sample and predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 65 and the maximum score was 422, for a range of 357. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 37, indicating a minimal potential for bias in outcome statistics. This is a reflection of the relatively high response rate (67%) in France, combined with an effective nonresponse adjustment steps carried out during weighting.

Data collection

Based on information provided on QC forms and during monthly QC conference calls, France generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Guidelines 8.1.1B and 8.1.2A on management of field staff.

France met a reduced requirement on training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. All of France's interviewers were provided with at least 15 hours of training.

France did not meet a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. France reached 10% for 100% of its interviewers. However, only completes were validated and not any other dispositions.

Instrument Data Quality

Translation

To the best of the Consortium's knowledge, France followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, France followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    a) Core items: 96.5%

    b) Literacy Items: 87.5%

    c) Numeracy Items: 92.3%

- Scoring reliability of paper-based national booklets
    d) Core items: 99.3%

    e) Literacy Items: 98.4%

    f) Numeracy Items: 98.8%

Assessment Data

Overall, 96.9% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In France, 71.5% of the respondents who completed the BQ took the computer-based cognitive assessment, while 26.3% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In France, 12.7% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 5.8% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding
To the best of the Consortium's knowledge, France followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labour Force Survey: Education: Standard met/Passed
- Comparison with Labour Force Survey: Occupation: Standard met/Passed
- Comparison with Labour Force Survey: Industry: Standard met/Passed

BQ Data
Background data were of very high quality for the France. If a respondent started the interview, the likelihood that she/he provided data is at a level above 99% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In France, about 90.5% of respondents reported income in exact amounts (88.6% across countries) and about 3.9% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In France, we observed 0.8% of cases with breakoffs.

Item Nonresponse
Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In France, these percentages were 18.6% for Literacy and 15.1% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in France, the percentage of nonresponse for Literacy was 10.6% and for Numeracy it was 7.8%. France did not administer the assessment for PSTRE.

Germany

Sampling

Germany followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. All QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
  - Home office: Through Consortium review of the preliminary SDIF, an anomaly was detected in the age distribution of the sample, with spikes at ages 30, 40, and 50. Germany investigated the reason for this pattern and discovered an error in the sample selection algorithm at the last stage of selection. Germany provided evidence that the sample remained probability-based despite this error and corrected the selection probabilities to reflect the actual selection algorithm used. However, they were unable to calculate exact selection probabilities, so the probabilities are based on a simulation.
  - In field: Not applicable
- Sample weighting: The Consortium followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights for Germany.
- Sampling error: Germany's design effect due to unequal weights is 1.22 for a sample size of 5,465. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 2,680. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (2.01). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. Germany's sample design involved an equal probability sample; however, the error in the sampling algorithm (see above) resulted in a variation in the selection probabilities. Further variation in the weights was added through nonresponse and calibration adjustments, although the Consortium followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 0.5% (undocumented immigrants).
  - Data collection: The weighted percentage of cases excluded because they are inaccessible was 2%.
- Weighted response rate: 55%
- Nonresponse bias analysis
  - Basic: Germany performed all required analyses. Its analysis showed significantly lower response propensities for age 26+, urban areas and non-Germans, based on registry information. Analysis of interviewer observation variables and area-level data from a consumer marketing survey also indicated lower response to PIAAC for lower education levels, lower socioeconomic status, higher rates of movers

and smaller household sizes. Age, municipality size, nationality, gender, region and education were used in weighting adjustments.

- Extended: Germany performed all required analyses. The extended analysis provides evidence that bias was reduced through the weighting adjustments.
  - Analysis 1 – Comparisons of estimates before and after weighting: The nonresponse bias in auxiliary variables (noted above in the Basic NRBA) was reduced through the weighting process. In addition, estimates of education and proxy proficiency changed substantially (relative difference > 2) as a result of the weighting adjustments. However, these estimates are not known for the full eligible sample, so it is difficult to make a conclusion about bias.
  - Analysis 2 – Comparisons of estimates to external totals: Significant differences were found between PIAAC estimates (using final weights) and Microcensus 2010 estimates of citizenship, municipality size, ISCED and work status. However, the estimates using the final weights are closer to the external totals than those using the base weights, with the differences diminished through weighting.
  - Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was slightly below average at 0.33 (0.30 for numeracy). The correlation between the raking dimensions and literacy scores was above average at 0.57 (0.58 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.61 (0.62 for numeracy), which was above the average across countries. Although Germany's response rate was low (55%), this analysis shows that weighting adjustments were effective in reducing NRB because of the correlation between the survey outcomes and the weighting variables. However, data users need to be cautioned that the analysis is based on correlations between the responding sample (55% of the selected sample) and the weighting variables. That is, the analysis assumes that same correlations exist for the remaining sampled cases that have no scores (45% of sampled cases).
  - Analysis 4 – Comparisons of estimates from alternative weightings: To calculate new weights, the final weighted data was repoststratified by each of the following variables separately: ISCED, citizenship, federal state and work status. The original weighting had used national education rather than ISCED, citizenship in nonresponse adjustment but not calibration, region but not federal state, and did not include work status. The mean literacy PV1 was significantly different when re-weighting by ISCED, but change was not substantial (differed by ~2). There were no other significant differences.
  - Analysis 5 – Analysis of variables collected during data collection: Germany looked at characteristics of the literacy-related nonrespondents and found that they belonged to the expected sociodemographic groups, providing evidence that this disposition code was used as intended. Bias was reduced by the LRNR weighting adjustment. Germany also reviewed

data from interviewer observation forms. The results confirmed its findings from the basic analysis.

- Analysis 6 – Level-of-effort analysis: Germany compared mean proxy proficiency scores, as well as education, work status and citizenship, between interviews conducted during the main release and interviews conducted during the second release of reissued cases. The mean proficiency score was significantly lower for high level-of-effort than low level-of-effort cases at the 10 percent significance level. There were no significant differences in the distribution of respondents' education, employment status, or citizenship status.
- Analysis 7 – Range of bias: The literacy scores' first plausible value was used to compute the range of scores within the responding sample and predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 78 and the maximum score was 406, for a range of 328. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 53, indicating a low potential for bias in outcome statistics. This is a reflection of an effective nonresponse adjustment strategy carried out during weighting. That is, even though Germany's response rate was low (55%), the effective nonresponse adjustment weighting reduced the potential bias in the outcome statistics to a low level. However, data users need to be cautioned that the analysis is based on assumptions about the range of proficiency scores for sampled cases that have no scores (45% of the sample).

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Germany generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 9.4.2 on interviewer training.

Germany met a reduced requirement on management. Guidelines 8.1.1B and 8.1.2A required weekly meetings between interviewers and supervisors and an interviewer-supervisor ratio of 20 or less. For the purpose of data evaluation, countries were considered to have met the standard if the meetings between interviewers and supervisors were held every other week and the interviewer-supervisor ratio was 30 or less. In Germany, interviewer-supervisor meetings occurred weekly and supervisor assignments ranged between 15 and 25 interviewers.

Germany partially met a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. The majority of Germany's validation cases were not selected randomly.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Germany followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Germany followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
  - Core items: 96.0%
  - Literacy Items: 97.9%
  - Numeracy Items: 95.8%
- Scoring reliability of paper-based national booklets
  - Core items: 99.9%
  - Literacy Items: 99.4%
  - Numeracy Items: 99.1%

Assessment data

Overall, 99.3% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Germany, 80.9% of the respondents who completed the BQ took the computer-based cognitive assessment, while 17.1% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Germany, 6.5% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 3.9% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Germany followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Germany. If a respondent started the interview, the likelihood that she/he provided data is at a level above 98% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Germany, about 90.9% of respondents reported income in exact amounts (88.6% across countries) and about 3.6% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Germany, we observed 1.5% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Germany, these percentages were 10.8% for Literacy and 7.6% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Germany, the percentage of nonresponse for Literacy was 7.0%, for Numeracy it was 3.8%, and for PSTRE it was 0.1%.

Ireland

Sampling

Ireland followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. Most QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
  - Home office: No issues
  - In field: No field issues detected
- Sample weighting: The Consortium followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights for Ireland.
- Sampling error: Ireland's design effect due to unequal weights is 1.37 for a sample size of 5,983. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 2,652. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (2.25). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. Ireland's sample design involved an unequal probability sample at the person level due to selecting one person no matter the household size. Further variation in the weights was added through nonresponse and calibration adjustments, although the Consortium followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 0.4% (The Geo-directory can underestimate mobile dwellings).
  - Data collection: N/A
- Weighted response rate: 72%
- Nonresponse bias analysis
  - Basic: Ireland performed all required analyses. Its analysis showed significantly lower response propensities in areas with lower levels of owner occupancy, areas with higher percentages of eligible non-Irish adults, areas where lower percentages of eligible adults spoke English as a native language, and areas with higher levels of unemployment. The overall response rate also varied by region (from 69% in Mid-East to 77% in South-West). There were no significant differences between respondents and nonrespondents across educational levels. Percentage non-English language spoken at home, percentage unemployment, percentage with lower secondary-level education or below, percent owner occupied, region, age, and gender were used in nonresponse adjustments.
  - Extended: Ireland performed analyses of comparison of weighted estimates to external totals, correlation of auxiliary variables and proficiency estimates and calculation of the range of potential bias. The preliminary extended analysis provides evidence that bias was reduced through the weighting adjustments.
    - Analysis 1 – Not required because overall response rate is above 70%.

- Analysis 2 – Comparisons of estimates to external totals: Differences were found between the PIAAC estimates (computed using final weights) and the 2011 census estimates of gender and educational attainment, but in percentage terms the overall shape of the distribution is very similar.
- Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was below average at 0.51 (0.51 for numeracy). The correlation between the raking dimensions and literacy scores was 0.50 (0.50 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.52 (0.53 for numeracy), which was about the average across countries. This indicates some potential for reducing NRB due to the correlation between survey outcome and weighting variables.
- Analysis 4 – Not required because overall response rate is above 70%.
- Analysis 5 – Not required because overall response rate is above 70%.
- Analysis 6 – Not required because overall response rate is above 70%.
- Analysis 7 – Range of bias: The literacy scores' first plausible value was used to compute the range of scores within the responding sample and predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 87 and the maximum score was 413, for a range of 326. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 27, indicating a minimal potential for bias in outcome statistics. This is a reflection of the very high response rate (72%) in Ireland. That is, even though the variables used for weighting had only moderate correlation with outcome scores, the high response rate has minimized the potential for nonresponse bias in the outcome statistics.

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Ireland appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 10.9.3 on fieldwork validation, Standard 9.4.2 on interviewer training and Guidelines 8.1.1B and 8.1.2A on management of field staff.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Ireland followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items.

All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Ireland followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
  - Core items: 97.1%
  - Literacy Items: 96.7%
  - Numeracy Items: 95.0%
- Scoring reliability of paper-based national booklets
  - Core items: 99.6%
  - Literacy Items: 99.2%
  - Numeracy Items: 99.3%

Assessment data

Overall, 97.6% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Ireland, 68.3% of the respondents who completed the BQ took the computer-based cognitive assessment, while 30.7% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Ireland, 19.4% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 4.3% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Ireland followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Ireland. If a respondent started the interview, the likelihood that she/he provided data is at a level above 99% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Ireland, about 90.8% of respondents reported income in exact amounts (88.6% across countries) and about 1.8% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Ireland, we observed 0.5% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Ireland, these percentages were 10.0% for Literacy and 7.5% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Ireland, the percentage of nonresponse for Literacy was 7.2%, for Numeracy it was 5.1%, and for PSTRE it was 0.1%.

Italy

Sampling

Italy followed the PIAAC technical standards and guidelines (TSG) related to sampling and weighting. All QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
  - o Home office: No issues
  - o In field: Not applicable
- Sample weighting: Italy followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights.
- Sampling error: Italy's design effect due to unequal weights is 1.43 for a sample size of 4,621. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 1,666. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (2.75). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. The goal of the sample design was to arrive at equal probabilities of selection for individuals. However, there was some variation observed in the base weights. Further variation in the weights was added through within-household sampling, nonresponse and calibration adjustments, although they followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population coverage
  - o Frame: The estimated percentage of the target population excluded from the frame was 0.8% (people in noninstitutional collective dwelling units).
  - o Data collection: The weighted percentage of cases excluded because they were inaccessible was 1.8%.
- Weighted response rate: 55%
- Nonresponse bias analysis
  - o Basic: Italy performed all required analyses. The required variables for education and employment were not available for use in the basic NRBA. The two-variable combination of age classes by gender, which was not used in weighting, showed some indications of potential nonresponse bias. Micro-regions, not used in weighting, did not show indications of potential bias. Indications of the potential for bias prior to weighting were found in age classes, household size, municipality size, and micro-region. Most significant specific categories are 16-25-year-olds (overrepresented) and 56-65 (under); 1 and 2 person households (under); large municipalities (under); North West (under) and North East and South (over). Among variables not used in weighting, the age by sex groups show possible underrepresentation for younger ages 16-34 for both sexes, and overrepresentation for 55-65 females. The logistic regression show significant effects among all six variables in the analysis.

- Extended (preliminary): Italy performed all required analyses. Its extended analysis provides evidence that bias was reduced through the weighting adjustments.
  - Analysis 1 – Comparisons of estimates before and after weighting: Significant differences 'before' NR adjustments among age classes, household size, and regions. Differences still over 2 standard errors away for single person household, North West, North East and South after NR adjustments, however, the standard errors (denominator) were small, which may overstate the size of the difference in the percentages. After calibration, in general the absolute differences were reduced for the regions, except for the South. Italy conducted significance testing that showed a slightly different picture, where significant potential bias remained for the North West only after the NR adjustment, among all the subgroups. Nonweighting variables were not used in the analysis.
  - Analysis 2 - Comparisons of estimates to external totals: Employment and education totals from the Labor Force Survey are significantly different from those from Italian Multipurpose Survey (used for PIAAC calibration). In order to explain these differences, it is important to note that the LFS is a rotated sample with the effect of attrition and substitution being allowed, while for the Italian Multipurpose Survey, the substitution is not allowed and is based on a two- stage sampling design of 60,000 units (observed sample persons). For Education, the largest absolute differences are for categories ISCED 3A-B and ISCED 2. For employment status, none of the confidence intervals for PIAAC and LFS overlap and the largest absolute differences are categories Not in Labor force and Unemployed. Italy provided an explanation of differences between PIAAC and the external source, and said it is not possible to say if these differences are due to a bias into the PIAAC estimates. For education, the largest differences (st_PIAAC - st_LFS) correspond to categories ISCED 3A-B and ISCED 2, where the relative differences are -8.0% and 5.5%, respectively. For employment status, the largest differences are for Not in Labor force and Unemployed (-2.6 % and 3.8%). For the reasons described above, it is not possible to say if these differences are due to a bias into the PIAAC estimates.
  - Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was below average at 0.22 (0.21 for numeracy). The correlation between the raking dimensions and literacy scores was average at 0.48 (0.52 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.49 (0.53 for numeracy), which was about the average across countries. This indicates some potential for reducing NRB due to the moderate correlation between the survey outcomes and the weighting variables. However, data users need to be cautioned that the analysis is based on correlations between the responding sample (55% of the selected sample) and the weighting

variables. That is, the analysis assumes that the same correlations exist for the remaining sampled cases that have no scores (45% of sampled cases).

- Analysis 4 – Comparisons of estimates from alternative weightings: The re-weighted estimate from the alternative more detailed education and employment status showed no important difference with the estimates based on the final weights. The overall difference is significant however, and with the alternative weights resulting in a higher average by four points. Therefore, there is some potential for bias in the resulting scores.
- Analysis 5 – Analysis of variables collected during data collection: The Italy weighting procedures separated the LRNR, therefore treating the LRNR cases appropriately. There were no domains with unexpected differences between LRNR and the comparison group. They provided frequencies from its NIR; however, only 133 completed the forms and therefore it is not possible to draw conclusions.
- Analysis 6 – Level-of-effort analysis: There were significant differences by level of effort for age class 46-55 (higher for low effort) and HH_size=1 (higher for low effort). This indicates that the thorough data collection efforts helped to reduce the bias due to nonresponse.
- Analysis 7 – Range of bias: The literacy scores' first plausible value was used to compute the range of scores within the responding sample and to predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 92 and the maximum score was 439, for a range of 347. Using weighting adjustments cells, and with an extreme assumption that nonrespondents would all score at the 10^{th} percentile within each weighting cell, and at the other extreme they would all score at the 90^{th} percentile within each weighting cell, the predicted maximum range of the mean was computed to be 62, indicating a low potential for bias in outcome statistics. This is a reflection of an effective nonresponse adjustment strategy carried out during weighting. That is, even though Italy's response rate was low (55%), the effective nonresponse adjustment weighting reduced the potential bias in the outcome statistics to a low level. However, data users need to be cautioned that the analysis is based on assumptions about the range of proficiency scores for sampled cases that have no scores (45% of the sample).

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Italy generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Guidelines 8.1.1B and 8.1.2A on management of field staff.

Italy met a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. Italy reached the 7% threshold for 99% of its interviewers.

Italy also met a reduced requirement on training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. Italian interviewers were provided with 27 hours of in-person training.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Italy followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Italy followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
  - Core items: 97.9%
  - Literacy Items: 97.0%
  - Numeracy Items: 96.2%
- Scoring reliability of paper-based national booklets
  - Core items: 99.4%
  - Literacy Items: 96.2%
  - Numeracy Items: 96.7%

Assessment data

Overall, 98.8% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Italy, 57.9% of the respondents who completed the BQ took the computer-based cognitive assessment, while 41.4% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Italy, 19.6% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 3.2% of those who reported having some

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Italy followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Italy. If a respondent started the interview, the likelihood that she/he provided data is at a level above 99% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Italy, about 80.3% of respondents reported income in exact amounts (88.6% across countries) and about 9.0% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Italy, we observed 0.7% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Italy, these percentages were 13.7% for Literacy and 10.3% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Italy, the percentage of nonresponse for Literacy was 12.8% and for Numeracy it was 9.0%. Italy did not administer the assessment for PSTRE.

Japan

Sampling

Japan followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. All QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
  - Home office: Japan had to adjust its sampling frame to take into account the unique circumstances caused by earthquake and tsunami. The alternative sample design deviates from an unbiased probability sample design as required by PIAAC's TSGs. However, an in-depth evaluation of the alternative approach indicated that the potential for bias in outcome statistics was expected to be minimal. Therefore, this alternative design is expected to produce national estimates for Japan that are comparable with other countries and with acceptable quality. Disproportionate sample allocation across strata Method was used in the alternative approach. Under this method, the maximum number of SPs allowed per PSU was 50. Strata with similar literacy levels were combined to reduce the impact on variances due to this upper bound limitation. This approach helped spread the sample across a larger number of PSUs, and reduce the sample weight variation. Initial base weights were adjusted to reflect all these changes.
  - In field: Not applicable
- Sample weighting: Japan followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create its weights.
- Sampling error: Japan's design effect due to unequal weights is 1.10 for a sample size of 5,278. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 3,362. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (1.54). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance.
- Japan started with an equal probability sample design. Due to changes (as described in sampling plan) the final design is an almost-equal probability sample. Further variation in the weights was added through nonresponse and calibration adjustments, although they followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 2.2% (non-nationals, undocumented immigrants).
  - Data collection: The weighted percentage of cases excluded because they are inaccessible was 2.8%.
- Weighted response rate: 50%
- Nonresponse bias analysis
  - Basic: Japan performed all required analyses.

- Extended: Japan performed all required analyses. The extended analysis provides evidence that bias was reduced through the weighting adjustments.
  - Analysis 1 – Comparisons of estimates before and after weighting: 19 variables were analyzed. Some of them are ratios, for example, the ratio of junior college or college graduate to the high school graduate. Estimates related to region, education, employment and age changed substantially (relative difference > 2) as a result of the weighting adjustments. However, these estimates are not known for the full eligible sample, so it is difficult to make a conclusion about bias. Half of the variables were used in weighting and the rest were not.
  - Analysis 2 - Comparisons of estimates to external totals: Japan took BQ variables to derive education and Labor force, and compared the estimates to the control totals. "PIAAC estimates were computed with final adjusted weights. Because the analysis variables are calibration variables, the estimates are just control totals."
  - Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was below average at 0.17 (0.20 for numeracy). The correlation between the raking dimensions and literacy scores was above average at 0.52 (0.51 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.53 (0.52 for numeracy), which was about the average across countries. Although Japan's response rate was low (50%), this analysis shows that weighting adjustments were moderately effective in reducing NRB because of the correlation between the survey outcomes and the weighting variables. However, data users need to be cautioned that the analysis is based on correlations between the responding sample (50% of the selected sample) and the weighting variables. That is, the analysis assumes that same correlations exist for the remaining sampled cases that have no scores (50% of sampled cases).
  - Analysis 4 – Comparisons of estimates from alternative weightings: To calculate new weights, the first plausible variable for literacy was used as a proficiency measure, and nonresponse adjustment cells were formed using different variables from those used in the original analysis, plus gender and age. Weights were calibrated using the same variables as in the original analysis. Proficiency estimates for respondents were obtained using the recalibrated weights. No differences were found in any domains.
  - Analysis 5 – Analysis of variables collected during data collection: Japan looked at type of building, floor in apartment building, and automatic lock house or apartment. People of higher socioeconomic class tend to occupy upper floors and live in automatic lock houses. The response rate of people living in apartments is low. Floor and automatic lock shown no differences.
  - Analysis 6 – Level-of-effort analysis: Japan compared number of visits, and developed a questionnaire to allow comparisons of response-related variables such as: at home vs. out, participate vs. refuse, cooperative vs.

annoying, and interested vs. not-interested. The first plausible value for literacy was used as a proficiency measure, and a regression analysis was performed. Japan's conclusion is as follows: High level-of-effort respondents due to not-at-home have such characteristics as male, young, and employed; its proficiency estimates are higher than those of stay-at-home respondents in every domain except not-in-Labor-force. Respondents who were cooperative and interested in the survey had such characteristics as young and highly educated; its proficiency estimates are higher than those of evasive respondents in every domain.

- Analysis 7 – Range of bias: The Literacy scores' first plausible value was used to compute the range of scores within the responding sample and to predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 126 and the maximum score was 418, for a range of 292. Using weighting adjustments cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 51, indicating a low potential for bias in outcome statistics. This is a reflection of an effective nonresponse adjustment strategy carried out during weighting. That is, even though Japan's response rate was low (50%), the effective nonresponse adjustment weighting reduced the potential bias in the outcome statistics to a low level. However, data users need to be cautioned that the analysis is based on assumptions about the range of proficiency scores for sampled cases that have no scores (50% of the sample).

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Japan appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Guidelines 8.1.1B and 8.1.2A on management of field staff.

Japan met a reduced requirement on training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. Japan provided 24 hours.

Japan partially met a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. Cases finalized as ineligible had no chance of being selected for validation, and the majority of validation cases were not selected randomly.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Japan followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Japan followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 99.2%
    - Literacy Items: 97.9%
    - Numeracy Items: 97.0%
- Scoring reliability of paper-based national booklets
    - Core items: 99.9%
    - Literacy Items: 99.8%
    - Numeracy Items: 99.7%

Assessment data

Overall, 99.9% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Japan, 61.8% of the respondents who completed the BQ took the computer-based cognitive assessment, while 36.8% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Japan, 17.9% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 12.1% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Japan followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Japan. If a respondent started the interview, the likelihood that she/he provided data is at a level above 98% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Japan, about 91.8% of respondents reported income in exact amounts (88.6% across countries) and about 3.0% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Japan, we observed 1.2% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Japan, these percentages were 6.5% for Literacy and 5.7% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Japan, the percentage of nonresponse for Literacy was 4.5%, for Numeracy it was 3.1%, and for PSTRE it was 0.0%.

Korea

Sampling

Korea followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. All QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
  - Home office: No issues
  - In field: No field issues detected
- Sample weighting: The Consortium followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights for Korea. Upon review of the distribution of raked weights, it was discovered that 20- to 26-year-olds were underrepresented in the sample and the raking dimension defined by age needed to be redefined to account for the underrepresentation. After discussions with Korea about this issue, a new raking dimension was submitted by Korea defined by age crossed with educational attainment.
- Sampling error: Korea's design effect due to unequal weights is 1.19 for a sample size of 6,667. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 5,086. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (1.31). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. The goal of the sample design was to arrive at equal probabilities of selection for households. However, there was some variation observed in the base weights. Further variation in the weights was added through within-household sampling, nonresponse and calibration adjustments, although they followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 2.4% (residents of small islands).
  - Data collection: Not applicable
- Weighted response rate: 75%
- Nonresponse bias analysis
  - Basic: Korea performed all required analyses. For the screener response rate, region, administrative district and residential type each showed statistical significance while region, residential type, gender, age, educational attainment, job type and household income were significantly different in the BQ response rate. Age, gender, occupation, urbanicity, region and education were used in weighting adjustments.
  - Extended: Not required

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Korea generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 10.9.3 on fieldwork validation, Standard 9.4.2 on interviewer training and Guidelines 8.1.1B and 8.1.2A on management of field staff.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Korea followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Korea followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 98.8%
    - Literacy Items: 99.1%
    - Numeracy Items: 96.7%
- Scoring reliability of paper-based national booklets
    - Core items: 100.0%
    - Literacy Items: 100.0%
    - Numeracy Items: 100.0%

Assessment data

Overall, 97.9% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Korea, 70.9% of the respondents who completed the BQ took the computer-based cognitive assessment, while 28.6% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Korea, 5.9% of respondents who reported having some computer experience

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.
[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

refused the CBA and took the PBA. An additional 10.4% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Korea followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Korea. If a respondent started the interview, the likelihood that she/he provided data is at a level above 99% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Korea, about 93.9% of respondents reported income in exact amounts (88.6% across countries) and about 1.5% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Korea, we observed 0.3% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Korea, these percentages were 7.4% for Literacy and 5.8% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Korea, the percentage of nonresponse for Literacy was 2.6%, for Numeracy it was 2.0%, and for PSTRE it was 0.2%.

The Netherlands

Sampling

The Netherlands followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. All QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
    - Home office: No issues
    - In field: Not applicable
- Sample weighting: The Netherlands followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create its weights.
- Sampling error: The Netherlands' design effect due to unequal weights is 1.10 for a sample size of 5,170. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 4,635. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (1.10). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. The Netherlands' sample design involved an equal probability sample. Variation in the weights was added through nonresponse and calibration adjustments, although they followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population coverage
    - Frame: The estimated percentage of the target population excluded from the frame was 0.9% (undocumented immigrants).
    - Data collection: The weighted percentage of cases excluded because they are inaccessible was 1.8%.
- Weighted response rate: 51%
- Nonresponse bias analysis
    - Basic: The Netherlands performed all required analyses. For all candidate auxiliary variables, except gender, the characteristics of the respondents and nonrespondents differ significantly. Therefore an inclusion of all candidate auxiliary variables, except gender, in the weighting model might result in a reduction of nonresponse bias.
    - Extended: The Netherlands performed all required analyses. The extended analysis provides evidence that bias was reduced through the weighting adjustments.
        - Analysis 1 – Comparisons of estimates before and after weighting were made for 19 variables. The variables included gender, age, generation, origin, degree of urbanization, group of provinces, household composition, social status, economic activity, type of dwelling, property-value of dwelling, monthly gross income, term of registration and low-, middle-, high-level of education. Estimates related to all the variables but gender

changed substantially (relative difference > 2) as a result of the weighting adjustments. Half of these variables were included in weighting.

- Analysis 2 - Comparisons of estimates to external totals: The Netherlands compared PIAAC estimates for education, employment status, occupation and industry to estimates from the Dutch Labor Survey. Statistical tests were not performed to check if differences are significant. Because the surveys differ in timing, observation mode, question wording, coding of education, profession and industry, performing proper statistical tests was found to be difficult. Therefore, it is hard to correctly interpret the differences in estimates.
- Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was below average at 0.26 (0.25 for numeracy). The correlation between the raking dimensions and literacy scores was above average at 0.57 (0.55 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.57 (0.55 for numeracy), which was above the average across countries. Although the response rate for The Netherlands was 51%, this analysis indicates potential for reducing NRB due to the high correlation between the survey outcomes and the weighting variables. However, data users need to be cautioned that the analysis is based on correlations between the responding sample (51% of the selected sample) and the weighting variables. That is, the analysis assumes that same correlations exist for the remaining sampled cases that have no scores (49% of sampled cases).
- Analysis 4 – Comparisons of estimates from alternative weightings: To calculate new weights, the first plausible variable for literacy variable was used as a proficiency measure, and the nonresponse adjusted weights were recalibrated using five of the original raking dimensions (gender by age, origin by generation, degree of urbanization by group of provinces, household composition, social status by income, term of registration in population registry), plus an alternative education variable. Proficiency estimates were obtained using the recalibrated weights. No differences were found.
- Analysis 5 – Analysis of variables collected during data collection: The Netherlands looked at the same 13 variables listed in Analysis 1 for the LR. Its conclusion states that "because some people of a first foreign background do not speak the Dutch language, it is considered not unlikely that 80% of the literacy related cases are people of a first generation foreign background". However, the Netherlands' weighting procedures separated the LRNR cases, therefore treating them appropriately.

2) An analysis of noninterview report data was not performed.

- Analysis 6 – Level-of-effort analysis: A variable with 3 levels was created: a person contacted 1-4 times, a person is contacted 5-6 times, and after initially refusing the person is contacted again. Two groups are formed: early respondents (1-4 contacts) and late respondents. A two-sample t-test was used to compare the literacy scores of these two groups.

Although the mean proficiency score of the late respondents is mostly higher than that of the other respondents, the differences are not significant.

- Analysis 7 – Range of bias: The literacy scores' first plausible value was used to compute the range of scores within the responding sample and to predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 81 and the maximum score was 440, for a range of 359. Using weighting adjustments cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 60, indicating a low potential for bias in outcome statistics. This is a reflection of an effective nonresponse adjustment strategy carried out during weighting. That is, even though The Netherlands' response rate was low (51%), the effective nonresponse adjustment weighting reduced the potential bias in the outcome statistics to a low level. However, data users need to be cautioned that the analysis is based on assumptions about the range of proficiency scores for sampled cases that have no scores (49% of the sample).

Data collection

The Netherlands partially met a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. The Netherlands reached the 7% threshold for 86% of its interviewers. Fourteen percent of interviewers were validated at less than the 7% level.

The Netherlands also partially met a reduced requirement on training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. About 60% of Netherlands's interviewers were provided with more than 15 hours; however, about 40% were provided with significantly fewer hours. The Netherlands offered significantly fewer training hours than recommended on all key aspects (gaining cooperation, BQ administration and assessment administration).

The Netherlands also partially met a reduced requirement on management. Guidelines 8.1.1B and 8.1.2A required weekly meetings between interviewers and supervisors and an interviewer-supervisor ratio of 20 or less. For the purpose of data evaluation, countries were considered to have met the standard if the meetings between interviewers and supervisors were held every other week and the interviewer-supervisor ratio was 30 or less. Netherlands' supervisor assignments were more numerous than the standard—55 interviewers, working on more than one project.

Instrument data quality

Translation

To the best of the Consortium's knowledge, the Netherlands followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, the Netherlands followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 95.6%
    - Literacy Items: 92.1%
    - Numeracy Items: 95.5%
- Scoring reliability of paper-based national booklets
    - Core items: 99.5%
    - Literacy Items: 99.9%
    - Numeracy Items: 99.9%

Assessment data

Overall, 98.2% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In the Netherlands, 87.5% of the respondents who completed the BQ took the computer-based cognitive assessment, while 9.7% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In the Netherlands, 4.5% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 3.2% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, the Netherlands followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for the Netherlands. If a respondent started the interview, the likelihood that she/he provided data is at a level above 97% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In the Netherlands, about 88.9% of respondents reported income in exact amounts (88.6% across countries) and about 4.5% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In the Netherlands, we observed 2.3% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In the Netherlands, these percentages were 10.0% for Literacy and 5.6% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in the Netherlands, the percentage of nonresponse for Literacy was 4.6%, for Numeracy it was 3.2%, and for PSTRE it was 0.2%.

Norway

Sampling

Norway followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. Most QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
  - Home office: No issues
  - In field: Not applicable
- Sample weighting: Norway followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights.
- Sampling error: Norway's design effect due to unequal weights is 1.05 with a sample size of 5,128. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 4,947. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (0.83). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. The effective sample size is set equal to the actual number of cases with plausible values since the overall design effect is less than 1.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 0.4% (undocumented immigrants).
  - Data collection: The weighted percentage of cases excluded because they are inaccessible was 0.4%.
- Weighted response rate: 62%
- Nonresponse bias analysis
  - Basic: Norway performed all required analyses. Chi-square analysis shows that there is significant dependence between response status and all the auxiliary variables except for gender, immigration category and country background. Age, special field, occupation, industry, income, region and education were used in BQ NR adjustments.
  - Extended: Norway performed most of the analysis except NIR (they do not have such data). The extended analysis provides evidence that bias was mostly reduced through the weighting adjustments
    - Analysis 1 – Comparisons of estimates before and after weighting: Most of the bias was reduced except for a few levels of certain categories, but there is no sign of significant bias (either rel diff<2 or abs diff<1 or both).
    - Analysis 2 - Comparisons of estimates to external totals: Significant differences were found between PIAAC estimates (using final weights) and different registers of the following categories: education (9), special field (2,6), occupation (2,4), and income after taxes (4). According to its reply, the number of people in the category "missing" tends to be

underestimated (for all variables), thus these people are probably underrepresented in our respondent sample. For other categories, the confidence interval contains the register total in most cases except for special field categories 2 and 6, occupation categories 2 and 4, and income category 4.

- Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was above average at 0.45. The correlation between the raking dimensions and literacy scores was below average at 0.23 (0.22 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.48, which was about average across countries. That is, weighting adjustments were not as effective in reducing bias, as compared to other countries, because of the level of correlation between the survey outcomes and the weighting variables. However, Norway had a higher than average response rate (62%), as compared to other countries, implying that the potential for bias could be somewhat lower as compared to countries with lower response rates.
- Analysis 4 –Comparisons of estimates from alternative weightings: Norway used 5-year age groups, immigration category, and income after taxes in the reweighting, and very little differences were found between the estimates using final weights and reweighted weights.
- Analysis 5 – Analysis of variables collected during data collection: Norway looked at characteristics of the literacy-related nonrespondents and found that they belonged to immigration groups, certain age groups (too young or too old), certain regions and lower education, providing evidence that this disposition code was used as intended. So bias was reduced by the LRNR weighting adjustment.
- Analysis 6 – Level-of-effort analysis: Norway defined level of effort by the interview time before or after December 31, 2011. There was significant difference in the distribution of respondents' education. The late respondents generally have a lower average proficiency score, except for immigration 3 group (Norwegian-born to immigrant parents), whose late respondents have a higher score than early ones'.
- Analysis 7 – Range of bias: The Literacy scores' first plausible value was used to compute the range of scores within the responding sample and to predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 67 and the maximum score was 441, for a range of 344. Using weighting adjustments cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 51, indicating a low potential for bias in outcome statistics. This is a reflection of the higher than average response rate (62%) in Norway. That is, as a result of achieving a relatively higher response rate, the potential for remaining bias is low even

though the weighting adjustments were not as effective, as compared to other countries, in reducing bias in outcome statistics.

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Norway generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 10.9.3 on fieldwork validation and Guidelines 8.1.1B and 8.1.2A on management of field staff.

Norway partially met a reduced requirement on training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. Overall training duration was significantly shorter than recommended. Some interviewers were offered significantly fewer training hours than recommended on all key aspects (gaining cooperation, BQ administration and assessment administration).

Instrument data quality

Translation

To the best of the Consortium's knowledge, Norway followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Norway followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 96.6%
    - Literacy Items: 96.5%
    - Numeracy Items: 95.9%
- Scoring reliability of paper-based national booklets
    - Core items: 99.0%
    - Literacy Items: 97.5%
    - Numeracy Items: 98.5%

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

Assessment data

Overall, 97.2% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Norway, 85.5% of the respondents who completed the BQ took the computer-based cognitive assessment, while 11.2% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Norway, 6.5% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 3.7% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Norway followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Norway. If a respondent started the interview, the likelihood that she/he provided data is at a level above 97% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Norway, about 97.6% of respondents reported income in exact amounts (88.6% across countries) and about 0.5% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Norway, we observed 2.2% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Norway, these percentages were 9.6% for Literacy and 8.1% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Norway, the percentage of nonresponse for Literacy was 5.2%, for Numeracy it was 3.6%, and for PSTRE it was 0.2%.

Poland

Sampling

Poland followed the technical standards and guidelines (TSG) related to sampling and weighting. All QC materials were completed fully and returned in a timely manner except some of the sample selection forms, which were not submitted until data collection started.

- Sampling Plan: No issues
- Sample selection
  - Home office: No issues
  - In field: Not applicable
- Sample weighting: The Consortium followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights for Poland. The only exception is that no separate adjustment for literacy-related nonrespondents (LRNR) was performed to avoid extreme weights, since none of the BQ LRNR have age and gender collected, and there is only one assessment LRNR.
- Sampling error: Poland's design effect due to unequal weights is 1.91 for a sample size of 9,366 adults ages 16-65. Poland oversampled 19-26-year-olds, which increases the design effect. Further variation in the weights was added through nonresponse and calibration adjustments, although the Consortium followed standard procedures to balance bias and variance. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 6,320. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (1.48). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 0.8% (undocumented immigrants and foreigners staying in Poland fewer than 3 months).
  - Data collection: The weighted percentage of cases excluded because they are inaccessible was 4.2%.
- Weighted response rate: 56%
- Nonresponse bias analysis
  - Basic: Poland performed all required analyses. Its analysis showed significantly lower response propensities for ages 26-35, areas with high education or low unemployment, and several regions. Age, gender, region, unemployment level, locality size, income level, number of cities, density of middle-school students, and density of middle-school students per computer with internet were used in weighting adjustments.
  - Extended: Poland performed all of the required analyses except the analysis on non-interview report form. The extended analysis provides evidence that bias was reduced through the weighting adjustments.

- Analysis 1 – Comparisons of estimates before and after weighting: Bias in age, area-level education, area-level unemployment, locality size, and region was reduced through the weighting process as most of these variables were used in weighting adjustments. Gender was also analyzed but it did not show bias between the respondent and eligible sample.
- Analysis 2 - Comparisons of estimates to external totals: PIAAC estimates (using final weights) are generally smaller than the Census 2011 estimates of age, gender, and region. Per Poland, "The registry and Census data were collected by two different institutions. Despite the fact there are significant differences between PIAAC estimates and external control totals, we have not made any adjustments because the relative frequencies of Age, Gender, and Region characteristics are virtually identical for Census and Registry data."
- Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was below average at 0.29 (0.28 for numeracy). The correlation between the raking dimensions and literacy scores was below average at 0.33 (0.30 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.40 (0.37 for numeracy), which was below the average across countries. That is, weighting adjustments were not as effective in reducing bias, as compared to other countries, because of the lower than average correlation between survey outcomes and weighting variables. Also data users need to be cautioned that the analysis is based on correlations between the responding sample (56% of the selected sample) and the weighting variables. That is, the analysis assumes that the same correlations exist for the remaining sampled cases that have no scores (44% of sampled cases).
- Analysis 4 – Comparisons of estimates from alternative weightings: To calculate new weights, the final weighted data were re-raked by employment status and education. The mean literacy scores by education are virtually the same before and after re-weighting. The mean literacy scores by employment status are slightly different before and after reweighting, which may be due to the random imputation of six cases with missing employment status.
- Analysis 5 – Analysis of variables collected during data collection: Poland looked at characteristics of the literacy-related nonrespondents and found that they belonged to the expected sociodemographic groups, providing evidence that this disposition code was used as intended. Poland did not perform analysis on data from non-interview report (NIR) forms. Per Poland, "In our opinion, NIR analysis does not bring any valid information since data were collected in open-ended form using two different kinds of software (TAO, CMS). There are significant differences between input data for the same respondents."
- Analysis 6 – Level-of-effort analysis: Poland does not have information on the number of contacts to define level-of-effort. So they compared interviews conducted in the first 6 months of data collection with the

interviews conducted in the last 3 months of data collection, assuming the interviews in the first 6 months required less effort than the last 3 months. The analysis variables include mean literacy scores, proportions by age, sex, area-level education, area-level unemployment, locality size, and region. Significant differences of mean literacy score were found between the two groups for some of the domains.

- Analysis 7 – Range of bias: The Literacy scores' first plausible value was used to compute the range of scores within the responding sample and to predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 59 and the maximum score was 446, for a range of 388. Using weighting adjustments cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 54, indicating a low potential for bias in outcome statistics. However, data users need to be cautioned that the analysis is based on assumptions about the range of proficiency scores for sampled cases that have no scores (44% of the sample).

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Poland generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 9.4.2 on interviewer training.

Poland met a reduced requirement on management. Guidelines 8.1.1B and 8.1.2A required weekly meetings between interviewers and supervisors and an interviewer-supervisor ratio of 20 or less. For the purpose of data evaluation, countries were considered to have met the standard if the meetings between interviewers and supervisors were held every other week and the interviewer-supervisor ratio was 30 or less. In Poland, meetings between supervisors and interviewers occurred only on an as-needed basis and/or biweekly.

Poland did not meet a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. Poland reached the 7% threshold for 40% of its interviewers. Sixty percent of interviewers were validated at less than the 7% level. Only some cases were selected randomly.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Poland followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items.

All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Poland followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 99.0%
    - Literacy Items: 97.3%
    - Numeracy Items: 96.0%
- Scoring reliability of paper-based national booklets
    - Core items: 99.6%
    - Literacy Items: 98.2%
    - Numeracy Items: 98.7%

Assessment data

Overall, 99.0% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Poland, 50.4% of the respondents who completed the BQ took the computer-based cognitive assessment, while 49.3% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Poland, 29.3% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 7.9% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Poland followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Poland. If a respondent started the interview, the likelihood that she/he provided data is at a level of 100% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Poland, about 81.8% of respondents reported income in exact amounts (88.6% across countries) and about 6.5% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Poland, we observed 0.0% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Poland, these percentages were 9.0% for Literacy and 6.2% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Poland, the percentage of nonresponse for Literacy was 8.4%, for Numeracy it was 5.3%, and for PSTRE it was 0.0%.

The Russian Federation[14]

Sampling

It is unclear whether the Russian Federation followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling due to the lack of information provided.

- Sampling Plan: During the sample design stage, the Consortium suggested increasing stratification levels and reducing the clustering to the maximum extent possible to reduce design effects. However, the Russian Federation was not able to implement this suggestion. The Russian Federation selected 25 PSUs (regions; three were self-representing) and 93 SSUs (cities, towns, villages).
- Sample Selection
    - Home office: The Russian Federation provided minimal information in their QC forms, so the Consortium was not able to adequately QC any stage of their sample selection.
    - In field: No field issues detected.
- Sample weighting: The Consortium followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights for the Russian Federation. A literacy-related nonresponse adjustment was not needed because there were no literacy-related non-respondents at any stage of the data collection. Also, BQ nonresponse adjustment was not conducted because the BQ response rate was close to 100%.
- Sampling error: The Russian Federation's design effect due to unequal weights is 2.09 for a sample size of 3 892. The Russian Federation's overall design effects are substantial due to the high level of clustering in the sample (small numbers of PSUs and SSUs), and thus failing to meet the quality measures (related to design effects) established for PIAAC. For example, the overall design effect for literacy is 15.77 (other Round 1 country design effects range from 0.80 to 3.81), and the effective sample size is 247 (the effective sample size for other Round 1 countries range from 1,666 to 7,848). The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, was computed as the number of cases with plausible values divided by the overall design effect for literacy. The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. Further variation in the weights was added through nonresponse and calibration adjustments, although the Consortium followed standard procedures to balance bias and variance.

Coverage and Nonresponse Bias

- Population Coverage

[14] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

- Frame: The estimated percentage of the target population excluded from the frame was 1.5% (Chechnya region, due to war in the region).
- Data collection: 1 220 cases were identified during the quality control processes that did not accurately reflect the true proficiencies of respondents. These respondents had implausible response times, duplicate cases, and abnormal response patterns and were excluded from the Russian Federation's database:

  1) *116 duplicate cases identified*. These were cases involving respondents with identical responses to items, response times, and number of actions to completion.

  2) *144 cases with an average response time per item less than 10 seconds*. Very rapid responses were best understood as being not representative of the respondents' skills.

  3) *949 cases collected by the most prolific 8 interviewers*. The cases collected by these interviewers were unusually homogeneous and had very different characteristics compared to other respondents in Russia.

  4) *11 cases from the same household with a sampled person that met one of the aforementioned criteria.* These cases were determined to be incongruent.

  The exclusion of data from the 8 most prolific interviewers resulted in the removal of all cases from the Moscow municipal area (two certainty PSUs). The final Russian data set is therefore representative of the Russian Federation resident population aged 16-65, excluding those residing in the Moscow municipal area[15].

- Weighted response rate: 52%
- Nonresponse bias analysis
  - Basic: The Russian Federation evaluated nonresponse bias at the screener stage only, because their self-computed BQ response rate was around 99%. The Russian Federation did not use all required analysis variables. Age, gender, education, and employment were not used because "virtually all refusals occurred at the initial stage of contact with any member of the household or the gatekeeper and it was not possible to obtain any information on household members." The Russian Federation performed all required analyses. Non-respondents tended to live in towns and villages, in regions with a higher percentage of employed people, and in regions with a lower percentage of people with higher education than respondents. The chi-square analysis also showed differential response rates by region, type of settlement (city, town, village), and level of education in the region.

[15] This approach was discussed with and validated by the PIAAC Technical Advisory Group (TAG). Copies of two memos prepared for the TAG outlining the criteria used to identify the cases for removal and the outcomes of this process are included at the end of this adjudication report.

- Extended: The Russian Federation did not perform all the required analyses using the final weights and proficiency scores. As a result, nonresponse bias could not be fully evaluated.
  - Analysis 1 – Comparisons of estimates before and after weighting: The Consortium was unable to determine whether bias in the auxiliary variables was reduced through the weighting process due to insufficient information provided for this analysis. The percentage distribution of sample cases at each weighting step at the screener level was not provided. Additionally, the definition of the eligible sample was unclear.
  - Analysis 2 – Comparisons of estimates to external totals: Differences were found between PIAAC estimates (using final weights) and census 2010 estimates of percent unemployed by region. In 13 of the 23 regions, the PIAAC unemployment rate was lower than that of the census estimate, which may be due to the possibility that unemployed or those who concealed their unemployment status categorically refused to take part in the survey, suggesting possible nonresponse bias.
  - Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the raking dimensions and literacy scores was 0.35 (0.34 for numeracy), which was below the average across countries. That is, weighting adjustments were not as effective in reducing bias, as compared to other countries, because of the low correlation between the survey outcomes and the weighting variables.
  - Analysis 4 – Comparisons of estimates from alternative weightings: This analysis was not performed using the final weights and proficiency scores.
  - Analysis 5 – Analysis of variables collected during data collection: This analysis was not conducted because there were no literacy-related nonrespondents.
  - Analysis 6 – Level-of-effort analysis: This analysis was not conducted due to the inability to classify respondents as difficult-to-contact. 99.6% of the respondents agreed to be interviewed after one follow-up attempt.
  - Analysis 7 – Range of bias: This analysis was not performed using the final weights and proficiency scores.

Data Collection

Based on information provided on QC forms and during monthly QC conference calls, the Russian Federation generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 9.4.2 on interviewer training and Standard 10.9.3 on fieldwork validation.

However, analysis of the data revealed evidence of a range of irregularities related to data collection (see above) affecting a significant proportion of cases, which should have been detected by validation. The fact that they were not detected suggests that validation was not conducted in a sufficiently rigorous manner. Therefore, the Russian Federation failed to meet the adjudication requirements on data collection validation.

The Russian Federation met a reduced requirement on management. Guidelines 8.1.1B and 8.1.2A required weekly meetings between interviewers and supervisors and an interviewer-supervisor ratio of 20 or less. For the purpose of data evaluation, countries were considered to have met the standard if the meetings between interviewers and supervisors were held every other week and the interviewer-supervisor ratio was 30 or less. The Russian Federation reported that meetings between their supervisors and interviewers occurred every other week.

Instrument Data Quality

Translation

To the best of the Consortium's knowledge, the Russian Federation followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for background questionnaire materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, the Russian Federation followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - g) Core items: 94.0%
    - h) Literacy Items: 86.7%
    - i) Numeracy Items: 91.5%
- Scoring reliability of paper-based national booklets
    - j) Core items: 100%

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

k) Literacy Items: 100%

l) Numeracy Items: 100%

Assessment Data

Overall, 99.1% of respondents who completed the background questionnaire (BQ) went on to take some cognitive assessment in either computer or paper format. In the Russian Federation, 66.5% of the respondents who completed the BQ took the computer-based cognitive assessment, while 33.4% took the paper-based assessment. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In the Russian Federation, 15.7% of respondents who reported having some computer experience refused the computer-based assessment and took the paper-based assessment. An additional 2.8% of those who reported having some computer experience failed the ICT Core and took the paper-based assessment. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the paper-based assessment.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, the Russian Federation followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

Background Questionnaire Data

Background data were of very high quality for the Russian Federation. If a respondent started the interview, the likelihood that she/he provided data is at a level of 100% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In the Russian Federation, about 81.6% of respondents reported income in exact amounts (88.6% across countries) and about 5.9% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In the Russian Federation, we observed 0.0% of cases with breakoffs.

Item Non-Response

3) Overall, the average proportions of non-response (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In the Russian Federation, these percentages were 11.6% for Literacy and 7.8% for Numeracy. Overall for computer-based items, the level of non-response was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in the Russian Federation, the percentage of non-response for Literacy was 12.2%, for Numeracy it was 7.3%, and for PSTRE it was 0.0%.

Data Adjudication Summary

As noted above analysis of the data from the Russian Federation revealed evidence of irregularities affecting a significant proportion of cases that were not picked up by validation. As a consequence, the Russian Federation failed to meet the adjudication requirements on data collection validation.

The TAG recommended and the OECD and Russian Federation agreed to remove from the database some 1 220 cases that were determined to not accurately reflect on the true proficiencies of respondents. Three criteria were used: very rapid response times, duplicate cases, and abnormal response patterns. Applying these criteria led to the exclusion of data from the 8 most prolific interviewers, which resulted in the removal of all cases from the Moscow municipal area. The remaining data met the minimum requirements for psychometric modelling and were subsequently scaled and weighted to represent the 16-65 year old population excluding residents of the Moscow region. .

The criteria for the removal of the most egregious cases relating to response time and duplication applied only to respondents assessed using CBA. Thus the number of potential cases not accurately representing the proficiency or the background variables of other respondents could not be fully evaluated. If such cases remain in the database it is likely we would observe reduced relationships between proficiency and background variables. For this reason, the TAG noted that, while the application of the three criteria would be likely to improve the fit and coherence of the data base, "*the deficiencies associated with the Russian data can neither be completely eradicated nor the accuracy of the data fully restored*". As a result, the Russian Federation received a Caution for Instrument Data Quality.

A Proposal to Improve Data Quality by
Filtering Incongruent Cases from the Most Recent Russian Database

Prepared by ETS, August 2013

Summary Statement unanimously supported by TAG:

The delivery of the Russian PIAAC data lacked timely evidential validity during sampling, data collection and database preparation which severely hampered the consortium's ability to validate the Russian data. In addition, a substantial number of cases were identified during the quality control processes that do not accurately reflect the true proficiencies of respondents. These respondents had implausible response times, duplicate cases, and aberrant response patterns. As a result, it was recommended and approved by the TAG and the OECD that these identifiable cases be dropped from the database. Moreover, it was recognized by the TAG that while the deficiencies associated with the Russian data can neither be completely eradicated nor the accuracy of the data fully restored, the removal of the three groups of respondents identified through the criteria suggested by the consortium will significantly improve the reliability and comparability of the Russian database.

Adjudication_RussianFederation_2013 09_DRAFT.DOCXContext

As requested by the OECD, ETS convened a virtual TAG meeting on July 29 to review and discuss quality issues surrounding the Russian data based on in depth analyses of the most recent database. There was unanimous agreement among the participating TAG members that the Russian data lacked sufficient quality with regards to reliability, validity and comparability. The major reasons for the poor overall quality of the data were identified as:

1) insufficient or untimely information provided to the consortium around sampling and survey operations undermined the evidential validity of the data;

2) a substantial number of respondents could not have read and answered the literacy and numeracy items correctly in the time-interval logged by the computer platform. More than 400 respondents were found to have an average time per item below 10 seconds, which is insufficient to process the sometimes extensive reading and stimulus material presented in the PIAAC tasks;

3) other quality control checks performed by the consortium resulted in the identification of a number of duplicated cases – 46 of these with responses and timing data that match exactly with other cases;

4) respondent data provided by the most prolific 8 interviewers are not consistent with other respondents in Russia; and,

5) a substantial number of items do not fit the common latent skill based psychometric model within the Russian data, and also do not fit the aggregate international database. These findings contradict the field test results, where such deviations were not observed, and also are incongruent with the PIAAC main test data from the Russian speaking sample from Estonia.

Rationale

Without knowing exactly what was done during sampling, survey administration, scoring and preparing the database, the inherent deficiencies observed in the data cannot be completely addressed without a full audit of the sampling and survey procedures. And, even if a full audit were feasible at this time it is unlikely that we will fully understand everything that has contributed to the incongruence in the Russian data.

However, it appears that the consortium can offer a proposal that will improve the quality of the Russian data by filtering out a significant proportion of the incongruent cases. Our suggestion is based on removing those cases that are identified as belonging to incongruent groups. These cases can be identified by applying a set of criteria that do not take into account the performance on the cognitive items. The consortium expects that applying these criteria will increase the coherence and comparability of the Russian national data as well as the fit of these data to the international database that forms the basis of the PIAAC.

Proposal for Salvaging the Majority of the Russian Data

While it has to be understood that procedure proposed below will not fully remediate the deficiencies present in the Russian database, it will provide the OECD with a strategy that helps to salvage more than 70% of the existing data. After careful analyses of the existing database, we suggest three criteria be applied to increase reliability and comparability of the Russian data. These include:

1) <u>Drop all duplicate cases that have been identified</u>. Duplicate cases involve respondents with exact same responses to items, the exact same response times, and the same number of actions. These duplicate cases are impossible to obtain without errors introduced by some form of intervention. These duplicated response patterns should be eliminated from the data as they do not represent the skills of two independent respondents.

2) <u>Drop all cases with an average response time per item less than 10 seconds</u>. It is nearly impossible to meaningfully respond to any open-ended questions involving multiple paragraphs in less than 10 seconds. Overall, the average response time for the Russian database is reduced compared to other countries participating in PIAAC (see Figures 1 and 2 below). We believe these very fast responses are best understood as being not representative of the respondents' skills. In part this understanding is based on our analyses of the reading components data with proficient readers and the time needed to respond to each of three components.

3) <u>Drop all cases collected by the "most prolific 8 interviewers"</u>. This group was identified not based on performance or other characteristics, but only on the fact that these

8 interviewers each provided many more cases than the other interviewers. Analyses show that these cases are unusually homogeneous and have very different characteristics when compared to other respondents in Russia. Their respondents produced nearly always correct answers on the majority of items and on some items nearly always incorrect responses (even below the level of the respondents interviewed by the remainder of interviewers, see figure 3). Their proportions correct do not resemble the rest of Russian data, and often contradict each other. They were nearly always incorrect on some of the easy items and nearly always correct on some of the very difficult items, which typically reflect erratic responding not related to the underlying skills. In contrast, most of the other respondents in the Russian database and in other countries show a systematic pattern between difficulty of the item and the skill of the respondent. These response patterns contribute to the poor fit of the measurement model and, therefore, do not represent true skills of respondents. In contrast, the respondents from the remaining interviewers show high congruence with the Russian speaking sample collected in Estonia (see figure 4), and exhibit a similar association when comparing these results in other pairs of countries (Figure 5).

In total, there are between 1400 and 1500 cases identified by these three criteria that should be dropped from the database because they are not representative of the true skills of respondents, and do not adequately reflect the distribution of the skills in the country. The removal of these cases together with proper weighting of remaining cases should increase the overall reliability and comparability of Russian data.

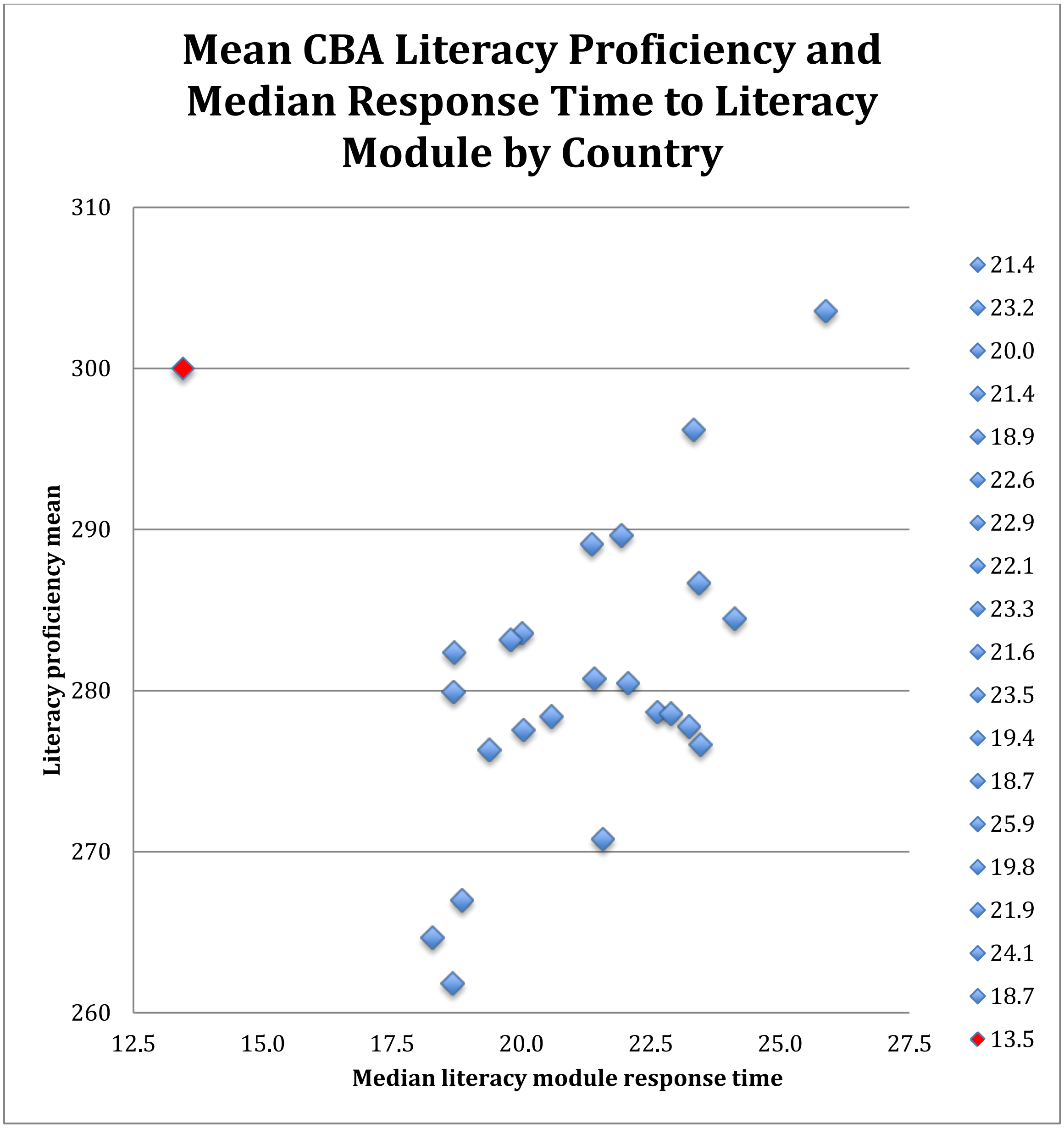


Figure 1: The unit of time is in minute for respondent who took Literacy module.

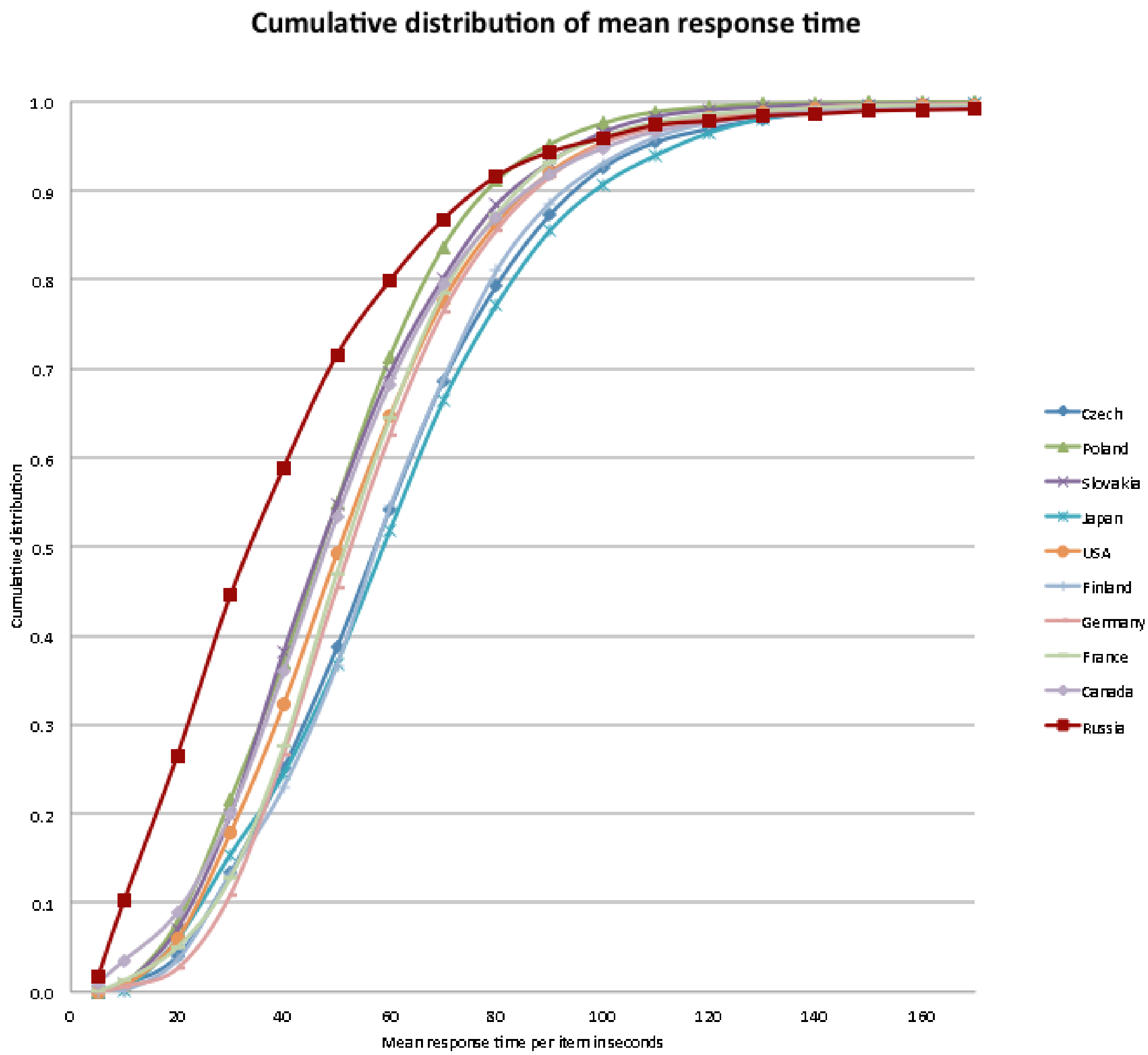


Figure 2: Mean item response time was calculated for those who took either Literacy or Numeracy or both core CBA items and/or CBA modules.

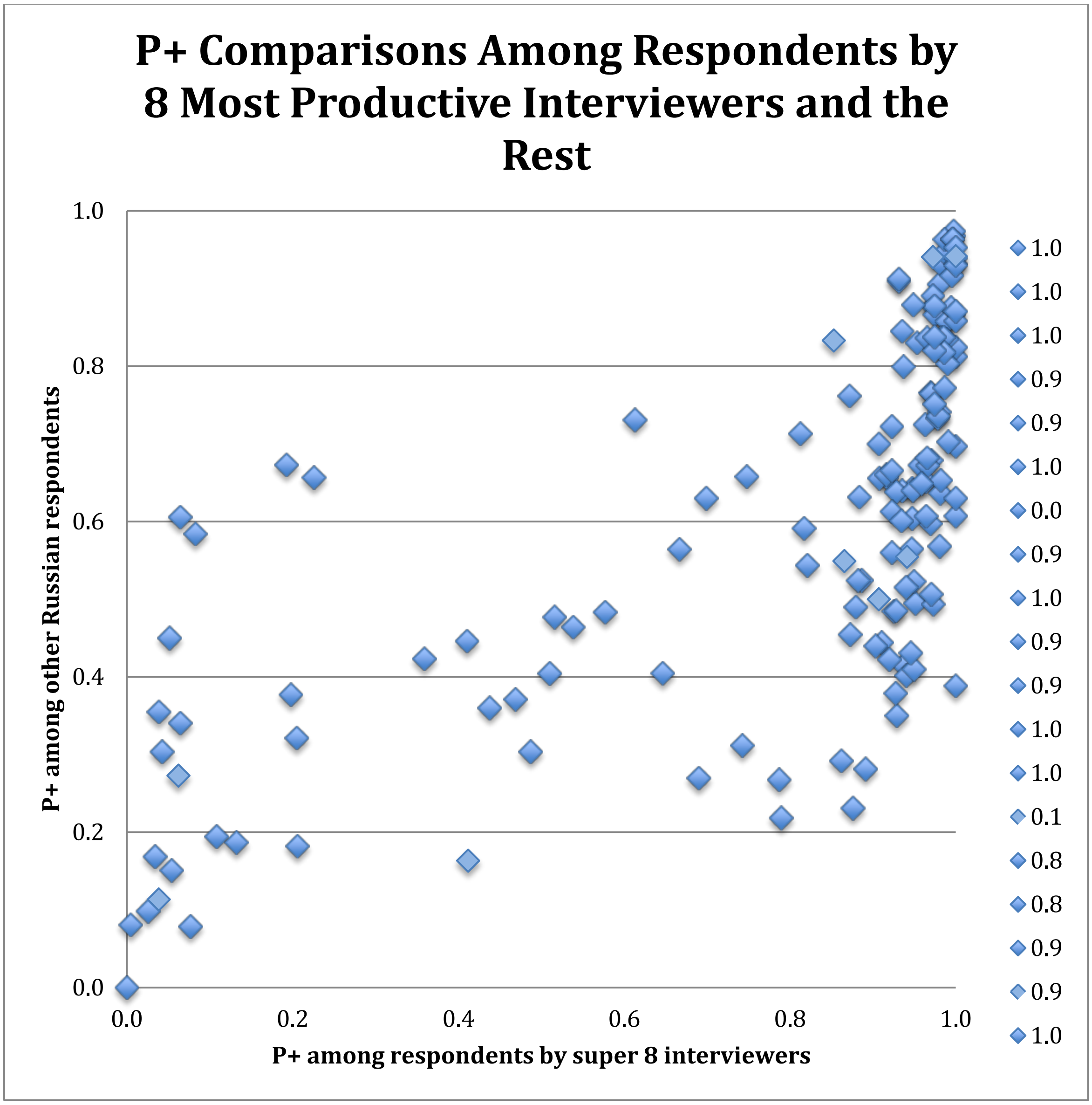


Figure 3: Eight out of 167 interviewers collected 1033 out of 5069 respondents. Two sets of P+ were calculated based on the 1033 cases and 4036 cases. Preliminary weights and standardized path weights for the CBA items were used. Above plots include both literacy and numeracy items. A very strong interaction of interviewers by P+ can be recognized that indicates that the data from the 8 most prolific interviewers does not align with the item P+ measures found in the remainder of the sample.

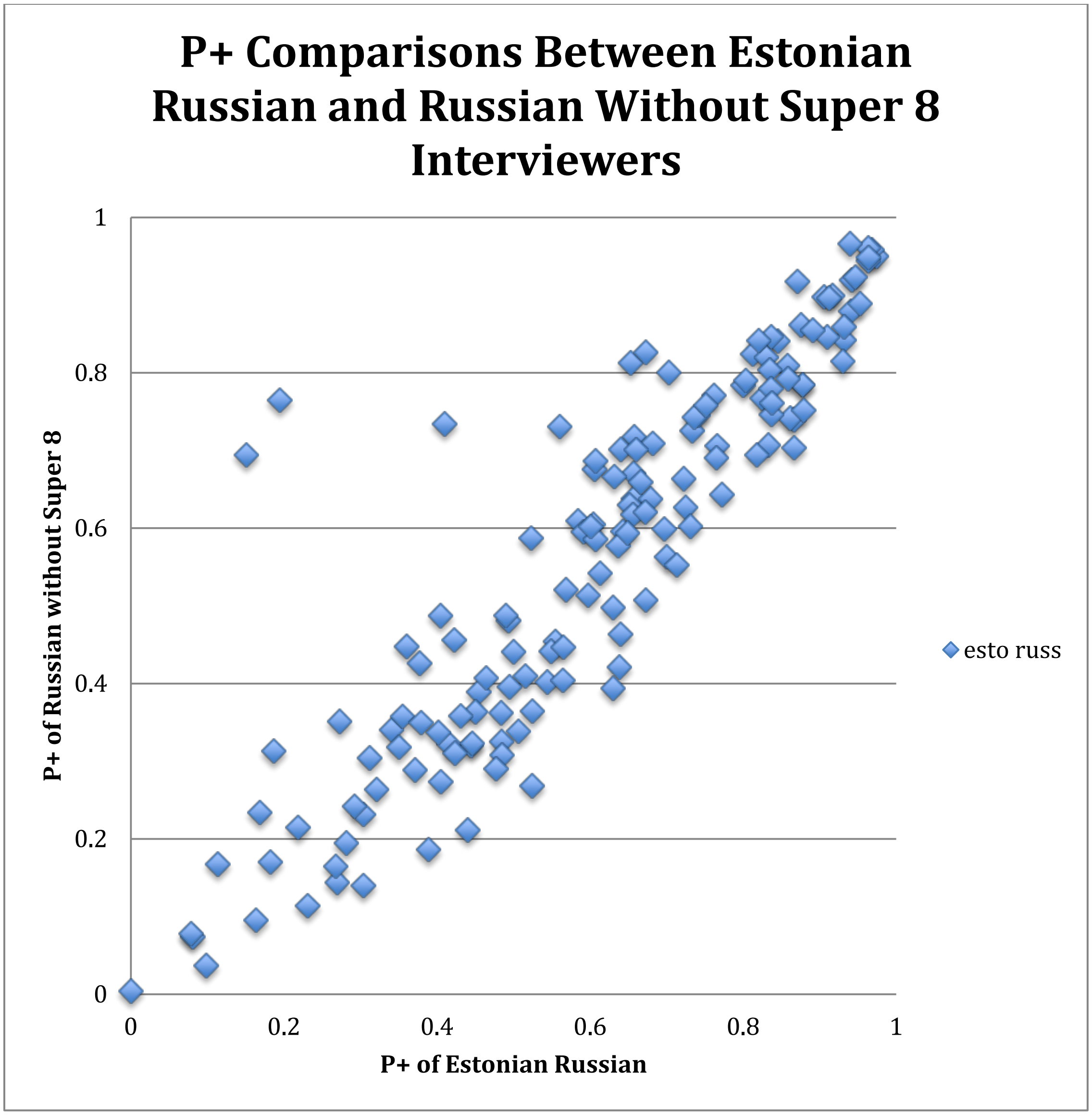


Figure 4: About 1400 of Russian speaking Estonians participated in the PIAAC survey in Estonia. The percent correct (P+) for this sample was compared against P+ of 4036 cases who were not associated with the 8 most prolific interviewers. The above plot includes both literacy and numeracy items. It can be seen that a very strong correlation of item P+ measures exists across the two samples.

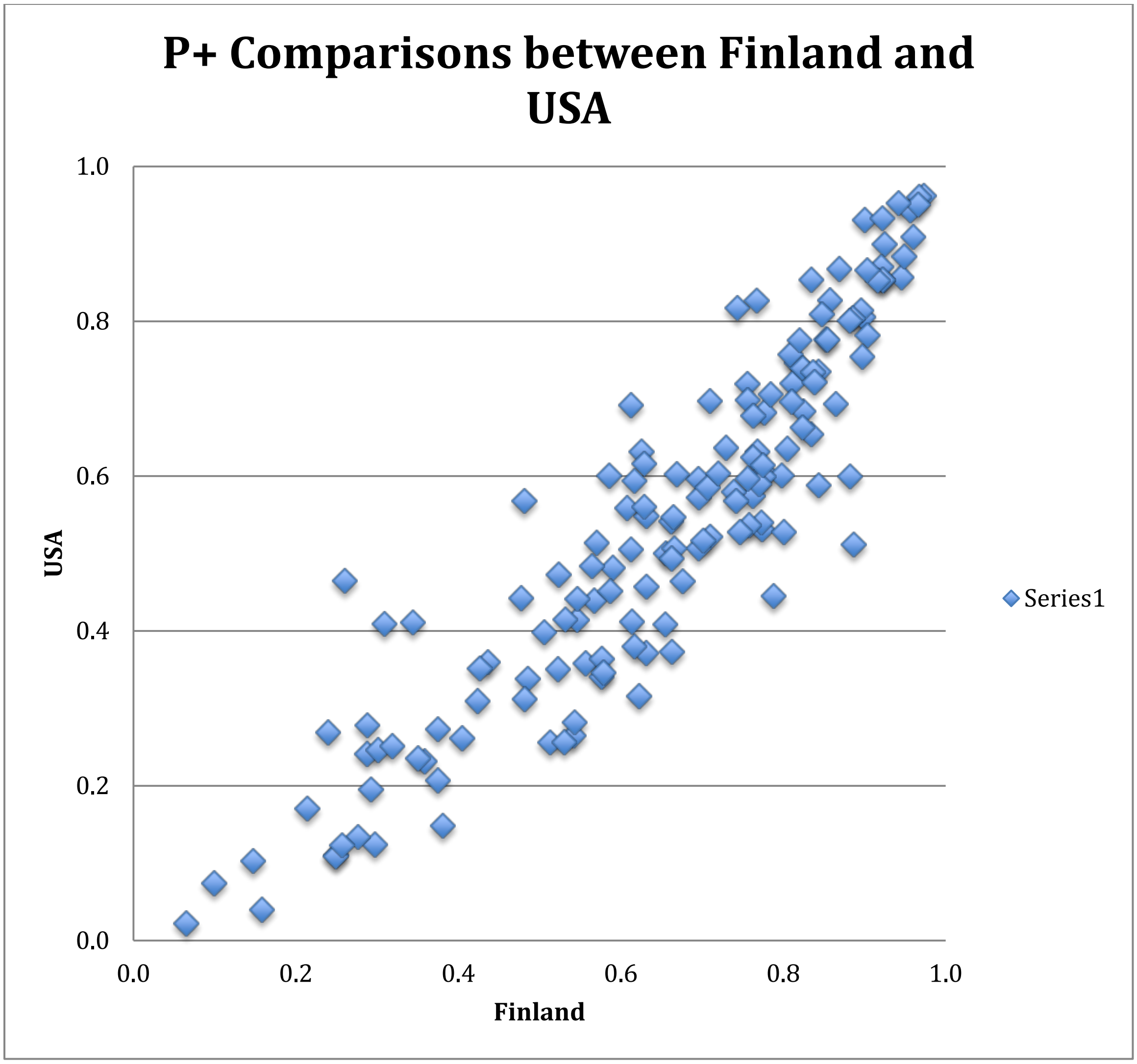


Figure 5: This plot is based on the P+ of Finland and USA and shows a very typical correlation of difficulty measures across countries. Despite of substantial differences in proficiency means of these two countries, the P+s have a very high correlation. The above plot includes both literacy and numeracy items.

A Note to the TAG regarding the outcomes of the process to improve the reliability, validity and comparability of data from the Russian Federation

OECD, September 2013

Introduction

At its meeting of July 29, the TAG reviewed the Russian data from PIAAC. At this meeting, the TAG concluded that:

> *The delivery of the Russian PIAAC data lacked timely evidential validity during sampling, data collection and database preparation, and severely hampered the consortium's ability to validate the Russian data. In addition, a substantial number of cases were identified during the quality control processes that do not accurately reflect the true proficiencies of respondents. These respondents had implausible response times, duplicate cases, and aberrant response patterns. As a result, it was recommended and approved by the TAG and the OECD that these identifiable cases be dropped from the database. Moreover, it was recognized by the TAG that while the deficiencies associated with the Russian data can neither be completely eradicated nor the accuracy of the data fully restored, the removal of the three groups of respondents identified through the criteria suggested by the consortium will significantly improve the reliability and comparability of the Russian database.*

The TAG helped establish this proposal to improve the fit of these data to the international database. This involved removing cases that were judged to belonging to incongruent groups from the data base. It is important to note that the groups in question were to be defined by applying a set of criteria that did not take into account performance on the cognitive items or the location where the interviewers collected the data. Removal of these groups of cases from the data base was expected to increase the coherence and comparability of the Russian national data as well as the fit of these data to the international database.

This proposal was implemented and the following groups of cases were excluded from the database:

> 1) *All duplicate cases identified*. These were cases involving respondents with identical responses to items, response times, and number of actions to completion.
>
> 2) *All cases with an average response time per item less than 10 seconds*. The 10 second criteria was chosen because it represents a set of cases with severely deviating response times; approximately 1/6th of the average response time per item observed for the other participating countries. Very rapid responses are best understood as being not representative of the respondents' skills (e.g. Wise & DeMars, 2005).

3) *All cases collected by the most prolific 8 interviewers*. The cases collected by these interviewers were unusually homogeneous and had very different characteristics compared to other respondents in Russia.

In total, some 1220 cases identified by the above three criteria were dropped from the Russian data base. The exclusion of the aberrant data from the 8 most prolific interviewers resulted in the removal of all cases that were identified by the Russian national PIAAC team as coming from the Moscow region. The final Russian data set is therefore representative of the Russian resident population aged 16-65, *excluding those residing in the Moscow metropolitan area*. The weighting procedures applied to the remaining cases assumed duplicated cases are random and rapid responders are not related to any of the background variables.

As a consequence of the removal of the aberrant cases, the fit of the Russian data to international item parameters was improved. The statistical properties of the sample showed more regularities, and the difficulties of the PIAAC items was more in line with the international sample, as well as with the sample taking the test in Russian collected as part of the population survey in Estonia.

Analysis of the resulting data indicates that the relationships between proficiency and the background variables usually associated with proficiency are considerably weaker in Russia than in other countries. However, the Russian PIAAC team has reviewed this and believes that it reflects particularities of the Russian society and economy.

The PIAAC adjudication process for the Russian data was reinitiated by first reviewing the cases that were dropped from the sample. Data collection validation (rechecks) is critical to data validity; it is the most important quality control feature of household data collection. Analysis of the dropped data revealed evidence that validation was not conducted in a manner that would detect possible falsification. Therefore, the Russian Federation failed to meet the adjudication requirements on data collection validation. Russia is the only country failing these requirements.

In addition, the Russian Federation sample failed to meet the PIAAC requirements for sample efficiency. The overall design effects are substantial mainly because the sample involves a high level of clustering in the sample. For example, the overall design effect for literacy is 15.77 (other Round 1 country design effects range from 0.80 to 3.81), and the effective sample size is 247 (the effective sample size for other Round 1 countries range from 1,666 to 7,848). The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, was computed as the number of cases with plausible values divided by the overall design effect for literacy. Russia is the only country failing these requirements.

The weighted response rate is equal to 52%. The correlation between weighting variables and outcome statistics was only 0.35 (other Round 1 countries correlations ranged from 0.37 to 0.70) indicating that weighting was not as effective in reducing bias as compared to other countries. However, the overall impact of nonresponse bias on the outcome statistics is unknown since the Russian Federation has not yet completed the required nonresponse bias analysis (as of September 20th).

Release of Russian Data

The OECD proposes to release the Russian data. Readers will be informed that the estimates for the Russian Federation relate to residents of the Russian Federation excluding Moscow in the following way:

> Results for the Russian Federation are included only in the data tables in the Annex to Chapter 2 of the report due to the timing of the processing of the Russian data.
>
> The data from the Russian Federation is *preliminary* and may be subject to change. Readers should note that the sample for the Russian Federation does not include the population of the Moscow municipal area. The data published, therefore, do not represent the entire resident population aged 16-65 years in Russia but rather the population of Russia *excluding* the population residing in the Moscow municipal area.
>
> More detailed information regarding the data from the Russian Federation as well as that of other countries can be found in the *Technical Report of the Survey of Adult Skills*.

Access to data from the Russian Federation will also be provided through the PIAAC Data Explorer and in in the form of a Public Use File.

The documentation provided in the Technical Report about the data from the Russian Federation will be the same as that provided for other countries. The public will have access to a full adjudication report covering compliance with the Technical Standards and Guidelines as well as information on process undertaken to improve the validity, reliability and comparability of the data as described above.

Members of the TAG are asked to:

- Establish that the recommendations from its meeting on 29 July have been appropriately implemented
- Agree that the note as stated above that will be included in the international report to qualify the data from the Russian Federation.

The Slovak Republic

Sampling

The Slovak Republic followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. All QC materials were completed fully.

- Sampling plan: No issues
- Sample selection
    - Home office: No issues
    - In field: Not applicable
- Sample weighting: The Consortium followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights for the Slovak Republic. An unknown eligibility adjustment was not needed because there were no inaccessible cases with unknown whereabouts.
- Sampling error: The Slovak Republic's design effect due to unequal weights is 1.23 for a sample size of 5,723. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 4,236. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (1.35). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. The Slovak Republic's sample design involved an equal probability sample. Variation in the weights was added through nonresponse and calibration adjustments, although the Consortium followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population coverage
    - Frame: The estimated percentage of the target population excluded from the frame was 0.07% (undocumented immigrants).
    - Data collection: The weighted percentage of cases excluded because they are inaccessible was 4.9%.
- Weighted response rate: 66%
- Nonresponse bias analysis
    - Basic: The Slovak Republic performed all required analyses. For required variables education and employment, PSU-level variables were used. The lowest weighted BQ response rate was in Bratislava region and other big cities (i.e., size of municipality more than 100,000 inhabitants). Moreover, females were more likely to respond as compared to males. PSUs with a lower employment rate and lower education degree achieved higher weighted BQ response rates. The classification tree analysis indicated that the response status was influenced by respondent's region, size of municipality, age cross gender, gender, and age category. Bratislavsky region had the highest nonresponse rate among all regions in the Slovak Republic (with higher ratio of middle-aged males). Large and medium-sized municipalities showed lower response-rate in comparison to small municipalities (except for Bratislavsky region). More nonrespondents were in the

middle category of persons aged 30-50 (seldom younger). The logistic regression showed significant relationships between response propensity and age, gender, region, size of municipality, employment, urbanicity, and education. All but education and employment were used in weighting.

- Extended: The Slovak Republic performed all required analyses, with questions pending on Analyses 4 and 6.
  - Analysis 1 – Comparisons of estimates before and after weighting: Bias in age, gender, region, municipality size, urbanicity, employment, and education was reduced through the weighting process.
  - Analysis 2 – Comparisons of estimates to external totals: The PIAAC estimates (calibrated using the Census 2011 control totals) of age, gender, region, and urbanicity were generally in line with the registry data. Some inconsistencies were found for the size of municipality. However, the Census data were deemed more reliable. It is the responsibility of each person to register with the local authorities when changing one's permanent or temporary residence, but this is rarely done in reality.
  - Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was below average at 0.33 (0.32 for numeracy). The correlation between the raking dimensions and literacy scores was below average at 0.33 (0.34 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.38 (0.38 for numeracy), which was below the average across countries. That is, weighting adjustments were not as effective in reducing bias, as compared to other countries, because of the low correlation between the survey outcomes and the weighting variables. However, Slovak had a relatively high response rate (66%), as compared to other countries, implying that the potential for bias could be lower in Slovak as compared to countries with lower response rates.
  - Analysis 4 – Comparisons of estimates from alternative weightings: To compute alternative weights, the final weighted data were recalibrated to registry data. Percentages (rather than proficiency estimates) were incorrectly provided for this analysis.
  - Analysis 5 – Analysis of variables collected during data collection: There were no significant differences between the literacy-related nonrespondents (n=22) and nonliteracy-related nonrespondents (n=5701) in terms of age, gender, region, size of municipality, urbanicity, employment, and education. To glean additional information on the nonrespondents, the Slovak Republic also examined its registry information and found that the highest proportions of nonrespondents were middle-aged males across all regions. Moreover, the ratio of older women aged 56-65 was higher in big cities compared to the same age category of males.
  - Analysis 6 – Level-of-effort analysis: The Slovak Republic defined level-of-effort by the number of visits required for the final disposition code that was obtained (early respondents were those needing two or less visits to

close the case, late respondents were those needing three or more visits to close the case). There were significant differences in the distribution of respondents' age, region, size of municipality, urbanicity, employment, and education. Since proficiency estimates were not provided, it is not possible to tell if high-level-of-effort respondents achieved higher or lower scores than low-level-of-effort respondents.

- Analysis 7 – Range of bias: The Literacy scores' first plausible value was used to compute the range of scores within the responding sample and to predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 97 and the maximum score was 390, for a range of 293. Using weighting adjustments cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 37, indicating a minimal potential for bias in outcome statistics. This is a reflection of the relatively high response rate (66%) in Slovak. That is, as a result of achieving a higher response rate, the potential for the remaining bias is low even though the weighting adjustments were not as effective, as compared to other countries, in reducing bias in outcome statistics.

Data collection

The Slovak Republic met a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. The Slovak Republic reached the 7% threshold for 97% of its interviewers.

The Slovak Republic also met a reduced requirement on training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. Slovak interviewers were provided with 20 hours of in-person training.

The Slovak Republic also met a reduced requirement on management. Guidelines 8.1.1B and 8.1.2A required weekly meetings between interviewers and supervisors and an interviewer-supervisor ratio of 20 or less. For the purpose of data evaluation, countries were considered to have met the standard if the meetings between interviewers and supervisors were held every other week and the interviewer-supervisor ratio was 30 or less. Meetings occurred every other week and supervisor assignments included 12 to 16 interviewers.

Instrument data quality

Translation

To the best of the Consortium's knowledge, the Slovak Republic followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, the Slovak Republic followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
  - Core items: 99.6%
  - Literacy Items: 95.0%
  - Numeracy Items: 96.1%
- Scoring reliability of paper-based national booklets
  - Core items: 100.0%
  - Literacy Items: 100.0%
  - Numeracy Items: 100.0%

Assessment data

Overall, 98.9% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In the Slovak Republic, 63.2% of the respondents who completed the BQ took the computer-based cognitive assessment, while 36.2% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In the Slovak Republic, 15.7% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 2.7% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, the Slovak Republic followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for the Slovak Republic. If a respondent started the interview, the likelihood that she/he provided data is at a level above 99% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In the Slovak Republic, about 84.5% of respondents reported income in exact amounts (88.6% across countries) and about 7.0% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In the Slovak Republic, we observed 0.3% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In the Slovak Republic, these percentages were 3.7% for Literacy and 3.3% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in the Slovak Republic, the percentage of nonresponse for Literacy was 5.4%, for Numeracy it was 3.5%, and for PSTRE it was 0.0%.

Spain

Sampling

Spain followed the PIAAC Technical Standards and Guidelines (TSG) related to sampling and weighting. Except for the End of data collection (SM-1) form that was not possible to generate, all QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
    - Home office: No issues
    - In field: Not applicable
- Sample weighting: Spain followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create its weights.
    - 3,266 of the 14,400 released cases were untraceable (disposition code 24 or 25).
    - Large variation in sample-person base weights (55.8529 - 12947.5). Spain needed to respect the minimum sample size required for each community.
- Sampling error: Spain's design effect due to unequal weights is 1.21 for a sample size of 6,055. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 4,710. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (1.27). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. Spain's sample design involved an unequal probability sample. Further variation in the weights was added through nonresponse and calibration adjustments, although they followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population coverage
    - Frame: Spain's frame did not have exclusions of the target population.
    - Data collection: The weighted percentage of cases excluded because they are inaccessible was 5%.
- Weighted response rate: 48%
- Nonresponse bias analysis
    - Basic: Spain performed all required analyses. Only base weights were used for all the analyses. Nonresponse is higher for age group 26-35, lower secondary level of education, nationality (ESP), and population in the third quartile of unemployment rate.
    - Extended: Spain performed all required analyses. The extended analysis provides evidence that bias was reduced through the weighting adjustments.
        - Analysis 1 – Comparisons of estimates before and after weighting: gender, age, degree of urbanization and employment rate showed reduction in bias through the weighting adjustments. These variables were used in weighting.
        - Analysis 2 - Comparisons of estimates to external totals: significant differences for "Full-time employed" and "other" were found between

PIAAC estimates (using final weights) and activity status from LFS (other categories: part-time employed, unemployed, pupil/student, apprentice/internship, retired/early retirement, permanently disabled, in compulsory military or community service, domestic work; no estimate, and therefore no comparison was done for the external source of apprentice/internship). This variable was not included in weighting. Definition is different in both surveys and it affects its comparison.

- Analysis 3 – Correlation of auxiliary variables and proficiency estimates: The correlation between the BQ nonresponse cells and literacy scores was above average at 0.53 (0.55 for numeracy). The correlation between the raking dimensions and literacy scores was above average at 0.59 (0.60 for numeracy). The correlation between literacy scores and the combination of nonresponse adjustment cells and raking dimensions was 0.62 (0.62 for numeracy), which was above the average across countries. Although the response rate for Spain was very low (48%), this analysis shows that weighting adjustments were effective in reducing NRB because of the high correlation between the survey outcomes and the weighting variables. However, data users need to be cautioned that the analysis is based on correlations between the responding sample (48% of the selected sample) and the weighting variables. That is, the analysis assumes that same correlations exist for the remaining sampled cases that have no scores (52% of the sampled cases).
- Analysis 4 – Comparisons of estimates from alternative weightings: To calculate new weights, Spain used the first plausible variable for literacy and numeracy as a proficiency measure, and re-raked the final weights using different categories of the same raking dimensions used in weighting (sex by age, and education by region) plus activity variable (described in Analysis 2). No differences were found.
- Analysis 5 – Analysis of variables collected during data collection: Spain compared the LR groups with the distribution of other nonrespondents. They looked at the variables used in weighting: age, gender, nationality, education, degree of urbanization and region. Spain's conclusion is that they "found significant differences in the variables considered. The groups in which the percentage of LR is greater than the comparison groups are: people over 56 years old, foreign people, and illiterate and Primary education levels." However, Spain's weighting procedures separated the LRNR cases, therefore treating them appropriately. Spain did not perform the non-interview report data.
- Analysis 6 – Level-of-effort analysis: Spain compared the number of attempts to contact a respondent. Two analyses were performed: a descriptive analysis of the number of attempts with the variables age and gender, and a regression analysis to compare the mean score of literacy and numeracy given the number of attempts (1-6 vs. more than 6). There are no significant differences between the groups. Among the completed cases, it has shown that six attempts were enough to get most of the respondents.

- Analysis 7 – Range of bias: The Literacy scores' first plausible value was used to compute the range of scores within the responding sample and to predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 64 and the maximum score was 394, for a range of 330. Using weighting adjustment cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 63, indicating a low potential for bias in outcome statistics. This is a reflection of an effective nonresponse adjustment strategy carried out during weighting. That is, even though Spain's response rate was very low (48%), the effective nonresponse adjustment weighting reduced the potential bias in the outcome statistics to a low level. However, data users need to be cautioned that the analysis is based on assumptions about the range of proficiency scores for sampled cases that have no scores (52% of the sample).
- Other – A significant test of the null hypothesis is that the probability of every dichotomous variable generated from the BQ variables does not depend on the nonresponse status. Variables included in the analysis were: highest education level and reading habits, from the LFS and Survey on Cultural Habits and Practices in Spain 2010-2011. These two surveys differ from PIAAC data collection and methodology, so results should be compared with caution. Most of the differences appear not to be significant; significant differences for "Full-time employed" and "In retirement or early retirement and other" agree with results obtained in Analysis 2. Spain's conclusion is that "the results suggest that nonresponse is not conditional on BQ variables."

Spain submitted an additional Extended Nonresponse Bias analysis (performed by Ricardo Mora from Universidad Carlos III Madrid). The analyses are different from those established by the consortium. Results show the same conclusions as the Extended NRBA conducted by the PIAAC team.

Data collection

Based on information provided on QC forms and during monthly QC conference calls, Spain generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 10.9.3 on fieldwork validation and Guidelines 8.1.1B and 8.1.2A on management of field staff.

Spain met a reduced requirement on interviewer training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. Spanish interviewers were provided with an average of 18 hours of in-person training.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Spain followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Spain followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
  - Core items: 97.7%
  - Literacy Items: 96.3%
  - Numeracy Items: 95.7%
- Scoring reliability of paper-based national booklets
  - Core items: 100.0%
  - Literacy Items: 99.9%
  - Numeracy Items: 100.0%

Assessment data

Overall, 97.3% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Spain, 66.0% of the respondents who completed the BQ took the computer-based cognitive assessment, while 33.1% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Spain, 13.0% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 7.1% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Spain followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Spain. If a respondent started the interview, the likelihood that she/he provided data is at a level above 99% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Spain, about 84.5% of respondents reported income in exact amounts (88.6% across countries) and about 4.4% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Spain, we observed 0.8% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Spain, these percentages were 14.5% for Literacy and 9.8% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Spain, the percentage of nonresponse for Literacy was 11.3% and for Numeracy it was 7.6%. Spain did not administer the assessment for PSTRE.

Sweden

Sampling

Sweden followed the technical standards and guidelines (TSG) related to sampling and weighting. All QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample Selection
  - Home office: No issues
  - In field: Not applicable
- Sample weighting: Sweden's weighting procedure is different from what is described in PIAAC Weighting and Variance Estimation plan. They did not conduct a separate adjustment for nonresponse and its unknown eligibility adjustment is the last step of weighting. However, its procedure adheres to the PIAAC standards.
- Sampling error: Sweden's design effect due to unequal weights is 1.13 for a sample size of 4,469. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 4,469. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (0.80). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. The effective sample size is set equal to the actual number of cases with plausible values since the overall design effect is less than 1.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was less than 1% (undocumented immigrants).
  - Data collection: The weighted percentage of cases excluded because they are inaccessible was 0%.
- Weighted response rate: 45%
- Nonresponse bias analysis
  - Basic: Sweden performed all required analyses. Its analysis showed significantly lower response propensities for people with low education, low income, not employed, age 26-35, certain occupations, and several regions. Education, region, employment, age, occupation, income, sex, country of birth, and year of immigration were used in weighting adjustments. Although the last three variables did not show significant relationship to response propensities, they were included in the weighting adjustment because it is known that they are related to proficiency and identify important subgroups.

- Extended: Sweden performed all of the required analyses except the 5th analysis below. The extended analysis provides evidence that bias was reduced through the weighting adjustments.
  - Analysis 1 – Comparisons of estimates before and after weighting: Bias in BQ education, employment status, and country of birth was reduced through the weighting process as similar register variables were used in weighting adjustments. Bias for employment benefits and social benefits was also reduced through the weighting process. Sweden also analyzed Skill use work – negotiating with people, Skill use everyday life – literacy –read books, literacy score, and numeracy score. It is hard to tell if bias was reduced for these variables since they are not available for nonrespondents.
  - Analysis 2 - Comparisons of estimates to external totals: PIAAC estimates (using final weights) were compared to both Labor Force Survey (LFS) 2011 and Census 2011 estimates. The differences between them for education, country of birth, region, occupation, and economic activity are in most cases not significant. There are some significant differences for employment status, probably caused by the different age coverage (LFS: 16-64- year- olds, Census: 15-64- year- olds) and definitions of employment status.
  - Analysis 3 – Correlation of auxiliary variables and proficiency estimates: Sweden did not perform a separate nonresponse adjustment in weighting. The correlation between literacy scores and the raking dimensions was 0.7 (0.7 for numeracy), which was the highest across countries. Although Sweden's response rate was very low (45%), this analysis shows that weighting adjustments were very effective in reducing NRB because of the high correlation between the survey outcomes and the weighting variables. However, data users need to be cautioned that the analysis is based on correlations between the responding sample (45% of the selected sample) and the weighting variables. That is, the analysis assumes that same correlations exist for the remaining sampled cases that have no scores (55% of sampled cases).
  - Analysis 4 – Comparisons of estimates from alternative weightings: The alternative weights were created by calibrating the weights using fewer and different cells (specifically, occupation and education by year of immigration were dropped from the calibration cell, and broader categories for country of birth were used). Sweden found only minor differences in the mean literacy score, distribution of education, employment status, and country of birth before and after re-weighting.
  - Analysis 5 – Analysis of variables collected during data collection: Sweden has not finished this analysis yet. Bias was reduced by the LRNR weighting adjustment.
  - Analysis 6 – Level-of-effort analysis: Sweden compared mean proficiency scores (both literacy and numeracy), as well as sex, age, education, employment status and country of birth, between low level-of-effort cases (1-3 contacts), medium level-of-effort cases (4-10 contacts) and high

level-of-effort cases (11+ contacts). There are no significant differences in the proficiency scores between easy, medium, and hard cases. There is a significant difference in the age group 56-65 years and people who are employed. People who are older or not employed are overrepresented among easy cases and underrepresented among hard cases. One might suspect that this would lead to lower proficiency score among easy cases than hard cases. There is no such effect though.

- Analysis 7 – Range of bias: The Literacy scores' first plausible value was used to compute the range of scores within the responding sample and to predict the range of estimates for nonrespondents. For the responding sample, the minimum score was 24 and the maximum score was 412, for a range of 389. Using weighting adjustments cells, and with an extreme assumption that nonrespondents would all score at the 10th percentile within each weighting cell, and at the other extreme they would all score at the 90th percentile within each weighting cell, the predicted maximum range of the mean was computed to be 54, indicating a low potential for bias in outcome statistics. This is a reflection of an effective nonresponse adjustment strategy carried out during weighting. That is, even though Sweden's response rate was very low (45%), the effective nonresponse adjustment weighting reduced the potential bias in the outcome statistics to a low level. However, data users need to be cautioned that the analysis is based on assumptions about the range of proficiency scores for sampled cases that have no scores (55% of the sample).

Data collection

Sweden partially met a reduced requirement on validation. Standard 10.9.3 called for the validation of 10% of cases for all (100%) interviewers, selected randomly across all dispositions. For the purpose of data evaluation, countries were considered to have met the standard if they had validated at least 7% of cases for at least 96% of its interviewers, selected randomly, across all dispositions. Sweden reached the 7% threshold for 91% of its interviewers. Nine percent of interviewers were validated at less than the 7% level.

Sweden also partially met a reduced requirement on training. For the purpose of data evaluation, countries were considered to have met the standard if they provided a minimum of 15 hours of training instead of the 30 hours required by the training programme provided by the Consortium. About half of Sweden's interviewers were provided with more than 15 hours; however, about half were provided with significantly fewer hours.

Sweden met a reduced requirement on management. Guidelines 8.1.1B and 8.1.2A required weekly meetings between interviewers and supervisors and an interviewer-supervisor ratio of 20 or less. For the purpose of data evaluation, countries were considered to have met the standard if the meetings between interviewers and supervisors were held every other week and the interviewer-supervisor ratio was 30 or less. Sweden's supervisor assignments included 23 interviewers.

Instrument data quality

Translation

To the best of the Consortium's knowledge, Sweden followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, Sweden followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

- Coding agreement of scoring anchor booklets
    - Core items: 96.5%
    - Literacy Items: 98.7%
    - Numeracy Items: 96.8%
- Scoring reliability of paper-based national booklets
    - Core items: 99.9%
    - Literacy Items: 99.8%
    - Numeracy Items: 99.9%

Assessment data

Overall, 96.9% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In Sweden, 90.1% of the respondents who completed the BQ took the computer-based cognitive assessment, while 9.4% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In Sweden, 5.2% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 2.8% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, Sweden followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for Sweden. If a respondent started the interview, the likelihood that she/he provided data is at a level of 100% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In Sweden, about 96.7% of respondents reported income in exact amounts (88.6% across countries) and about 1.1% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In Sweden, we observed 0.0% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In Sweden, these percentages were 13.5% for Literacy and 9.1% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in Sweden, the percentage of nonresponse for Literacy was 6.2%, for Numeracy it was 4.3%, and for PSTRE it was 0.3%.

The United States

Sampling

The United States followed the technical standards and guidelines (TSG) related to sampling and weighting. All QC materials were completed fully and returned in a timely manner.

- Sampling plan: No issues
- Sample selection
  - Home office: No issues
  - In field: Not applicable
- Sample weighting: The United States followed the procedures in the PIAAC Weighting and Variance Estimation Plan to create weights.
- Sampling error: The United States' DEFF due to unequal weights is 1.27 for a sample size of 5,010. The effective sample size, which is the sample size needed to achieve the same sampling variance as a simple random sample, is 2,211. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy (2.21). The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. The United States' sample design involved an equal probability sample of dwelling units. Further variation in the weights was added through within-household sampling, nonresponse and calibration adjustments, although the United States followed standard procedures to balance bias and variance.

Coverage and nonresponse bias

- Population coverage
  - Frame: The estimated percentage of the target population excluded from the frame was 0.08% (people in a gated community).
  - Data collection: The weighted percentage of cases excluded because they were inaccessible was 0%.
- Weighted response rate: 70%.
- Nonresponse bias analysis
  - Basic: The United States performed all required analyses. At the screener level, only the percentage of the population below 150% of the poverty level was significant. This indicates that there are fewer nonrespondents in the higher poverty levels. At the BQ level, the NRBA found the following variables that were significant at the α=0.05 level: region; percent of the population below 150% of the poverty level; percent of the population age 18-64 that is employed; age category; indicator for children under age 16 in household; and gender. The multivariate analysis identified the lowest response rate for the following characteristics:
    - Hispanics age 26 and older,
    - With no children in the household,
    - Not living in the Northeastern United States,
    - Living in segments with unemployment exceeding 4.8 percent, and

- Living in areas (Census tracts) with less than 5.1 percent of the population being linguistically isolated.

The presence of children in the household was a dominant variable in distinguishing response rate groups. In general, younger persons were found to be more available to participate in an in-person household survey, as are those with children ages 16 and younger, and women.

- Since all significant variables in both the screener and BQ analyses were used in the respective weighting adjustments, the potential for nonresponse bias should be reduced by those adjustments.
- One source of undercoverage was the portion of the population that does not have a usual home. This is primarily the homeless population. An attempt was made to correct this minor level of noncoverage (estimated to be less than 1%) by including poverty indicators in the nonresponse adjustment. The only other known undercoverage of the population was in a particular segment in the Western region that was selected for the survey but to which our survey staff were not granted access.

- Extended: The analysis was not required since the weighted response rate was greater than or equal to 70%.

Data collection

Based on information provided on QC forms, the United States generally appears to have met the original requirements as described in the PIAAC Technical Standards and Guidelines (TSG), in particular Standard 10.9.3 on fieldwork validation, Standard 9.4.2 on interviewer training and Guidelines 8.1.1B and 8.1.2A on management of field staff.

Instrument data quality

Translation

To the best of the Consortium's knowledge, the United States followed the PIAAC Technical Standards and Guidelines (TSG) associated with translation and verification, in particular, Standard 6.1 for new cognitive items, Standard 6.2 for BQ materials, and Standard 6.3 on linking cognitive items. All adaptations were documented and all materials went through full verification[1] prior to the Field Test and a partial verification[2] prior to the Main Survey.

- Outcome: TSG followed/Passed

Scoring

To the best of the Consortium's knowledge, the United States followed the PIAAC Technical Standards and Guidelines (TSG) associated with scoring paper-and-pencil instruments, in particular, Standard 11.3.

---

[1] Full verification was a sentence by sentence check for equivalence to source + linguistic correctness + appropriate/approved adaptations, with a final check that crucial issues identified during verification have been correctly addressed in pre-final instruments.

[2] Partial verification is a check of correct echoing of FT to MS changes in source version + vetting and verification of other changes at the initiative of countries, with again a final check on crucial.

- Coding agreement of scoring anchor booklets
    - Core items: 99.1%
    - Literacy Items: 99.5%
    - Numeracy Items: 97.3%
- Scoring reliability of paper-based national booklets
    - Core items: 99.1%
    - Literacy Items: 97.2%
    - Numeracy Items: 98.9%

Assessment data

Overall, 98.9% of respondents who completed the BQ went on to take some cognitive assessment in either computer or paper format. In the United States, 79.9% of the respondents who completed the BQ took the computer-based cognitive assessment, while 14.9% took the PBA. Across all countries, 73.5% of respondents who completed the BQ took the computer-based form of the assessment and 23.9% took the paper-based form.

Some respondents who reported having computer experience refused to take the PIAAC assessment in computer-based format. Thus, these respondents took the paper-based form of the assessment. In the United States, 6.9% of respondents who reported having some computer experience refused the CBA and took the PBA. An additional 4.3% of those who reported having some computer experience failed the ICT Core and took the PBA. Overall, across all countries, 11.8% of respondents who reported computer experience refused to take the assessment on the computer and 4.7% failed the ICT Core and were therefore routed to the PBA.

The captured data for reading components showed no anomalies in terms of accuracy and missing data. Recorded time showed similar characteristics from what was seen in the Field Test in relationship to the skill of respondents.

The assignment of cognitive modules within the Virtual Machine accurately followed the intended workflow. That is to say, the administration of Literacy, Numeracy and PSTRE modules followed the assessment design and the adaptive routing within the Literacy and Numeracy modules was accurately implemented. Analysis also showed accurate data capture for all countries.

Coding

To the best of the Consortium's knowledge, the United States followed the PIAAC Technical Standards and Guidelines (TSG) associated with coding, in particular, Standard 11.2.

- Double coding Occupation: Standard met/Passed
- Double coding Industry: Standard met/Passed
- Comparison with Labor Force Survey: Education: Standard met/Passed
- Comparison with Labor Force Survey: Occupation: Standard met/Passed
- Comparison with Labor Force Survey: Industry: Standard met/Passed

BQ data

Background data were of very high quality for the United States. If a respondent started the interview, the likelihood that she/he provided data is at a level above 95% with practically only one exception: Income related questions are reported either in exact monetary amounts or in broad categories. In the United States, about 93.4% of respondents reported income in exact amounts (88.6% across countries) and about 1.5% reported income in broad categories (4.2% across countries). If a respondent decided to break off the interview, the interviewer was able to collect a reason for the breakoff. The data contains about 2.0% cases with breakoff codes across countries, which indicate that the reason for breakoffs were either language related issues, reading writing issues, or disabilities. In the United States, we observed 4.2% of cases with breakoffs.

Item nonresponse

Overall, the average proportions of nonresponse (omitted or not reached) for the paper-based items were 10.8% for Literacy and 7.6% for Numeracy. In the United States, these percentages were 12.3% for Literacy and 6.5% for Numeracy. Overall for computer-based items, the level of nonresponse was 7.2% for Literacy, 4.9% for Numeracy, and 0.1% for PSTRE. For computer-based items in the United States, the percentage of nonresponse for Literacy was 5.3%, for Numeracy it was 3.7%, and for PSTRE it was 0.2%.

Table 7F-4. PIAAC Data Quality Evaluation Table – Sampling

Country	Sampling: Sample Design and Selection: Sampling Plan	Sampling: Sample Design and Selection: Sample Selection: Home Office	Sampling: Sample Design and Selection: Sample Selection: In Field	Sampling: Sample Weighting	Sampling: Sampling Error (DEF): Unequal Weighting Effect	Sampling: Sampling Error (DEF): Effective Sample Size[1]
	(1.A)	(1.B)	(1.C)	(1.D)	(1.E)	
Australia	P	C-U	C-U	C-PC	1.6	3,061
Austria	P	P	NA	P	1.09	3,561
Canada	P	P	P	P	2.76	7,848
Cyprus[2]	P	P	P	P	1.39	2,855
Czech Republic	P	C-NC	P	P	2.88	1,725
Denmark	P	P	NA	P	1.27	5,861
England (UK)	P	P	C-PC	P	1.35	2,176
Estonia	P	P	NA	P	1.04	3,785
Finland	P	P	NA	P	1.05	5,464
Flanders (Belgium)	P	P	NA	P	1.04	3,215
France	P	P	NA	P	1.05	6,867
Germany	P	C	NA	P	1.22	2,680
Ireland	P	P	P	P	1.37	2,652
Italy[3]	P	P	P	P	1.43	1,666

P: Pass (relevant requirement completely met)
C: Caution (relevant requirement met to a reasonable extent)
C-A: Caution, approved deviation
C-NC: Caution, did not comply
C-PC: Caution, partial compliance
C-U: Caution, quality level unknown due to country confidentiality restrictions or unavailability of data
F: Fail

[1]The effective sample size is the sample size needed to achieve the same sampling variance as a simple random sample. The effective sample size was computed as the number of cases with plausible values divided by the overall design effect for literacy. The overall design effect incorporates the design effects due to sampling variance (unequal weights, stratification and clustering) and imputation variance. The effective sample size is set equal to the actual number of cases with plausible values for countries where the overall design effect is less than or equal to 1.

[2] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.

Table 7F-4 (cont). PIAAC Data Quality Evaluation Table – Sampling

Country	Sampling					
	Sample Design and Selection			Sample Weighting	Sampling Error (DEF)	
	Sampling Plan	Sample Selection			Unequal Weighting Effect	Effective Sample Size[1]
		Home Office	In Field			
	(1.A)	(1.B)	(1.C)	(1.D)	(1.E)	
Japan	P	C-A	NA	P	1.1	3,362
Korea	P	P	P	P	1.19	5,086
Netherlands	P	P	NA	P	1.1	4,635
Northern Ireland (UK)	P	P	C-PC	P	1.54	563
Norway	P	P	NA	P	1.05	4,947
Poland	P	P	NA	C	1.9	6,320
Russian Federation [3]	C-PC	C-NC	P	P	2.09	247
Slovak Republic	P	P	NA	P	1.23	4,236
Spain	P	P	NA	P	1.21	4,710
Sweden	P	P	NA	P	1.13	4,469
United States	P	P	P	P	1.27	2,211

P: Pass (relevant requirement completely met)
C: Caution (relevant requirement met to a reasonable extent)
C-A: Caution, approved deviation
C-NC: Caution, did not comply
C-PC: Caution, partial compliance
C-U: Caution, quality level unknown due to country confidentiality restrictions or unavailability of data
F: Fail

[3] Please refer to the note regarding the Russian Federation in the Note to Readers section of this report.

Table 7F-5. PIAAC Data Quality Evaluation Table – Coverage and Nonresponse Bias)

Country	Coverage and Nonresponse Bias											
	Population Coverage (Exclusions)		Weighted Reponses Rate (RR) and Coverage Rate (CR)		Nonresponse Bias Analysis (NRBA)							
						Extended[1]						
	Frame	Data Collection	Weighted RR	Weighted CR	Basic	1	2	3	4	5	6	7
	(2.A)	(2.B)	(2.C)		(2.D)	(2.E)						
Australia	3.30%	NA	71%	69%	P	NA	NA	NA	NA	NA	NA	NA
Austria	0.60%	0.80%	53%	52%	P^U	P^U	P	P^- 2	P	P	P^+	P^- 1
Canada	1.80%	NA	59%	57%	P	P	P	P^- 3	P	P	P^+	P^- 1
Cyprus[2]	<2.0%	NA	73%	72%	P	NA	P^C	P^- 3	NA	P	P^+	NA
Czech Republic	1.80%	NA	66%	65%	P	P^C	P^C	P 2	C-U	P	P^+	P 2
Denmark	<0.1%	5.00%	50%	48%	P	P	P^C	P^- 3	P	P	P^+	P^- 4
England (UK)	2.00%	NA	59%	58%	P	P	P^C	P^- 3	C-NC	P^U	C-U	P^- 1
Estonia	2.80%	0.60%	63%	61%	P	P^U	P	P 4	P	P	P^+	P 1
Finland	0.20%	0.50%	66%	66%	P^U	P	P	P 2	P	P	P	P 1
Flanders (Belgium)	1.00%	4.00%	62%	59%	P^U	P^U	P^C	P 4	P	P	P^+	P 1
France	<2.6%	2.40%	67%	63%	P^U	P	C-NC	P 2	C-NC	C-NC	C-NC	P 1
Germany	0.50%	2.00%	55%	54%	P	P	P	P^- 2	P	P	P	P^- 2
Ireland	0.40%	NA	72%	72%	P	NA	P^C	P 3	NA	NA	NA	P 2
Italy	0.8%[3]	1.90%	56%	54%	P^U	P^U	P	P^- 3	P^c	P	P^+	P^- 4
Japan	2.20%	2.80%	50%	47%	P	P	P^U	P^- 3	P	P	P^+	P^- 2

P: Pass (relevant requirement completely met)
C: Caution (relevant requirement met to a reasonable extent)
C-A: Caution, approved deviation
C-NC: Caution, did not comply
C-PC: Caution, partial compliance
C-U: Caution, quality level unknown due to country confidentiality restrictions or unavailability of data
F: Fail

[1] See explanation on page following the end of this table.
[2] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[3] Italy's population exclusions was estimated to be 0.8%, however, the estimate does not include the illegal immigrant population. No estimate of the percentage of illegal immigrant population was available.

Table 7F-5. PIAAC Data Quality Evaluation Table – Coverage and Nonresponse Bias

Country	Coverage and Nonresponse Bias											
	Population Coverage (Exclusions)		Weighted Reponses Rate (RR) and Coverage Rate (CR)		Nonresponse Bias Analysis (NRBA)							
						Extended[1]						
	Frame	Data Collection	Weighted RR	Weighted CR	Basic	1	2	3	4	5	6	7
	(2.A)	(2.B)	(2.C)		(2.D)	(2.E)						
Korea	2.40%	NA	75%	73%	P	NA	NA	NA	NA	NA	NA	NA
Netherlands	0.90%	1.80%	51%	50%	P	P	P	P^{-} 2	P	P	P^{+}	P^{-} 2
Northern Ireland (UK)	2.00%	NA	65%	64%	P	P	P^{C}	P 2	C-NC	P^{U}	C-U	P 1
Norway	0.40%	0.40%	62%	62%	P	P	P^{C}	P 4	P	P	P^{+}	P 2
Poland	0.80%	4.20%	56%	53%	P	P	P^{C}	P^{-} 4	P	P	P^{+}	P^{-} 2
Russian Federation [4]	1.50%	NA	52%	51%	P^{U}	C-PC	P^{C}	P^{-} 4	C-NC	NA	C-U	C-NC
Slovak Republic	0.10%	4.90%	66%	63%	P	P	P	P 4	C-PC	P	C-PC	P 1
Spain	0.00%	5.00%	48%	46%	P	P^{U}	P	C 2	P	P	P^{+}	C 2
Sweden	<1.0%	0.00%	45%	45%	P	P	P	C 1	P	P^{U}	P^{+}	C 2
United States	0.10%	NA	70%	70%	P	NA	NA	NA	NA	NA	NA	P

P: Pass (relevant requirement completely met)
C: Caution (relevant requirement met to a reasonable extent)
C-A: Caution, approved deviation
C-NC: Caution, did not comply
C-PC: Caution, partial compliance
C-U: Caution, quality level unknown due to country confidentiality restrictions or unavailability of data
F: Fail

[4] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.

Table 7F-5 (cont). PIAAC Data Quality Evaluation Table – Coverage and Nonresponse Bias

Explanation regarding footnote No. 1:

There are four types of Pass: P = Pass, P+ = Pass with evidence of bias reduction (used for Analysis 6 only), PU = Pass with only a partially completed analysis (i.e., the quality level is unknown) due to unavailability of data, PC = Pass with caution because there are some indications of some signification differences without further explanation, leading to a possible indication for some limited potential for bias, PIR = Pass with only one item with item response rate below 85%.

For Analysis 3, the codes represent the following:

P:	RR	>= 60%	Moderate
P-:	RR	50-60%	Low
C:	RR	< 50%	Very low

1:	Correlation >=.65	Very High
2:	.55 <= Correlation < .6 5	High
3:	.45 <= Correlation < .55	Moderate
4:	.35 <= Correlation < .45	Low
5:	Correlation < .35	Very low

For Analysis 7, the codes represent the following:

P:	RR	>=60%	Moderate
P-:	RR	50-60%	Low
C:	RR	< 50%	Very low

1:	Range of Bias < 50	Minimal
2:	50 <= Range of Bias < 65	Low
3:	65 <= Range of Bias < 80	Moderate
4:	80 <= Range of Bias < 95	High
5:	Range of Bias >= 95	Very High

Table 7F-6. PIAAC Data Quality Evaluation Table – Data Collection and Instrument Data Quality

Country	Data Collection			Instrument Data Quality					
	Field Validation / Back-checks	Training	Management	Translation	Scoring	Assessment data	Coding	BQ Data	Item Nonresponse BQ[1]
	(3.A)	(3.B)	(3.C)	(4.C)	(4.D)	(4.A)	(4.D)	(4.B)	(4.E)
Australia	P	P	P	P	P	P	P	P	P
Austria	C-PC	P	C-PC	P	P	P	P	P	P
Canada	C-PC	P	P	P	P	P	P	P	P
Cyprus[2]	P	P	P	P	P	P	P	P	P
Czech Republic	P	C-PC	C-PC	P	P	P	P	P	P[IR]
Denmark	P	P	P	P	P	P	P	P	P
England (UK)	C-NC[8]	C-PC	P	P	P	P	P	P	P
Estonia	P	P	P	P	P	P	P	P	P[IR]
Finland	C-NC[3]	P	P	P	P	P	P	P	P
Flanders (Belgium)	C-PC	P	P	P	P	P	P	P	P
France	C-NC[4]	P	P	P	P	P	P	P	P
Germany	C-PC	P	P	P	P	P	P	P	P
Ireland	P	P	P	P	P	P	P	P	P
Italy	P	P	P	P	P	P	P	P	P[IR]
Japan	C-PC	P	P	P	P	P	P	P	P

P: Pass (relevant requirement completely met)
C: Caution (relevant requirement met to a reasonable extent)
C-A: Caution, approved deviation
C-NC: Caution, did not comply
C-PC: Caution, partial compliance
C-U: Caution, quality level unknown due to country confidentiality restrictions or unavailability of data
F: Fail

[1] See explanation on page following the end of this section.
[2] Please refer to notes A and B regarding Cyprus in the *Note to Readers* section of this report.
[3] 7% or more for 46% FIs; less than 7% from 54% FIs
[4] Only completes were validated

Table 7F-6. PIAAC Data Quality Evaluation Table – Data Collection and Instrument Data Quality

Country	Data Collection			Instrument Data Quality					
	Field Validation / Back-checks	Training	Management	Translation	Scoring	Assessment data	Coding	BQ Data	Item Nonresponse BQ[3]
	(3.A)	(3.B)	(3.C)	(4.C)	(4.D)	(4.A)	(4.D)	(4.B)	(4.E)
Korea	P	P	P	P	P	P	P	P	P
Netherlands	C-PC	C-PC	C-PC	P	P	P	P	P	P
Northern Ireland (UK)	C-PC	C-PC	P	P	P	P	P	P	P
Norway	P	C-PC	P	P	P	P	P	P	P
Poland	C-NC[6]	P	P	P	P	P		P	P[IR]
Russian Federation[5]	F[7]	P	P	P	P	C[9]	P	P	P
Slovak Republic	P	P	P	P	P	P	P	P	P[IR]
Spain	P	P	P	P	P	P	P	P	P
Sweden	C-PC	C-PC	P	P	P	P	P	P	P
United States	P	P	P	P	P	P	P	P	P

P: Pass (relevant requirement completely met)
C: Caution (relevant requirement met to a reasonable extent)
C-A: Caution, approved deviation
C-NC: Caution, did not comply
C-PC: Caution, partial compliance
C-U: Caution, quality level unknown due to country confidentiality restrictions or unavailability of data
F: Fail

[5] Please refer to the note regarding the Russian Federation in the *Note to Readers* section of this report.
[6] 7% or more for 40% FIs; less than 7% for 60% FIs
[7] Based on information provided on quality-control (QC) forms and during monthly QC conference calls, the Russian Federation followed validation requirements. However, analysis of the data revealed irregularities affecting a significant proportion of cases. This level of irregularities should have been detected by validation. The fact that it was not suggests that validation was not conducted in a manner sufficiently adequate to detect it. Therefore, the Russian Federation did not meet the requirements on validation.
[8] 7% or more for 20% FIs; less than 7% for 80% FIs
[9] See "Data Adjudication Summary" section in the Russian Federation Adjudication Report for details.

Table 7F-6. PIAAC Data Quality Evaluation Table – Data Collection and Instrument Data Quality

Explanation regarding footnote No. 1:
There are four types of Pass: P = Pass, P+ = Pass with evidence of bias reduction (used for Analysis 6 only), PU = Pass with only a partially completed analysis (i.e., the quality level is unknown) due to unavailability of data, PC = Pass with caution because there are some indications of some signification differences without further explanation, leading to a possible indication for some limited potential for bias, PIR = Pass with only one item with item response rate below 85%.

For Analysis 3, the codes represent the following:

P:	RR	>= 60%	Moderate
P-:	RR	50-60%	Low
C:	RR	< 50%	Very low

1:	Correlation >=.65	Very High
2:	.55 <= Correlation < .6 5	High
3:	.45 <= Correlation < .55	Moderate
4:	.35 <= Correlation < .45	Low
5:	Correlation < .35	Very low

For Analysis 7, the codes represent the following:

P:	RR	>=60%	Moderate
P-:	RR	50-60%	Low
C:	RR	< 50%	Very low

1:	Range of Bias < 50	Minimal
2:	50 <= Range of Bias < 65	Low
3:	65 <= Range of Bias < 80	Moderate
4:	80 <= Range of Bias < 95	High
5:	Range of Bias >= 95	Very High