Cross-Cultural Personality Research:

Conceptual and Methodological Issues

Verónica Benet-Martínez

University of California at Riverside

To appear in R. W. Robins, C. Fraley, R. F. Krueger (Editors), *Handbook of Research Methods in Personality Psychology*. Gilford Press.

Running title: Cross-Cultural Personality Research
Keywords: personality, cross-cultural, cultural, methods, research
Date: June 2006

Corresponding Author:
Verónica Benet-Martínez
Department of Psychology
University of California Riverside
Riverside, CA 92521
email: veronbm@ucr.edu

CONTENTS

I. Introduction

Personality is shaped by both genetic and environmental factors; among the most important of the latter are cultural influences. Culture consists of shared meaning systems that provide the standards for perceiving, believing, evaluating, communicating, and acting among those who share a language, a historic period, and a geographic location (Triandis, 1996). More recently Chiu and Chen (2004) have defined culture as "a network of knowledge that is produced, distributed, and reproduced among a collection of interconnected people" (p. 173). Culture is transmitted through language, media messages, cultural practices and institutions, and through the modeling of behavior. Cultural influences on personhood were a prevalent concern in early personality psychology (e.g., Allport, 1954; Kluckhohn & Murray, 1948; McClelland, 1961), but largely ignored in modern personality theory and research until the early 1990s. However, many cultural studies conducted during the last decade on issues such as self-processes, emotion, and personality traits have firmly established the following: culture is a key determinant of what it means to be a person (see reviews by Church, 2000; Diener, Oishi, & Lucas, 2003; Triandis & Suh, 2002).[1] Further, our personality –the affective, motivational, and cognitive dispositions that influence our evaluations and reactions to the environment– cannot be separated from the broad social and cultural context where it develops and is expressed. In fact, as Markus and Kitayama (1998) eloquently say:

> A cultural psychology approach assumes that personality … is completely interdependent with the meanings and practices of particular sociocultural contexts. People develop their personalities over time through their active participation in the various social worlds in which they engage. A cultural psychological perspective implies that *there is no personality without culture*; there is only a biological entity. (p. 67; italics added).

Most personality psychologists would agree that the systematic study of how culture influences social and intra-personality behavior should be an essential part of our discipline. Yet, cultural studies continue to be somewhat underrepresented in personality psychology. Why is this? One reason may be historical. Because of the serious methodological, theoretical, and

ethical limitations of several cultural anthropological studies conducted in the middle decades of the 20th century (e.g., 'culture and personality' or 'national character' studies), some psychologists may still view culture studies of personality with skepticism (LeVine, 2001). In addition, cultural psychological studies pose unique methodological challenges (concerning the conditions and parameters of research) to traditional personality psychology. Comparisons across cultural groups are often challenging and expensive; many cultural studies require painstaking translations, new hypotheses and instruments, expensive overseas trips, and networking with foreign researchers who are familiar with the cultures under study. Further, several factors may limit the interpretation of results, including problems in translation, the presence of response biases, and the unfamiliarity of respondents in some cultures with the use of rating scales.

Despite the above challenges, cultural research offers scientists exciting and interesting benefits and opportunities not available with traditional research approaches (Matsumoto, 2000). Cultural personality studies help elucidate how macro contextual factors mediate and moderate personality outcomes (e.g., McCrae, 2001; Schimmak et al., 2002), help dispel shaky cultural stereotypes (e.g., Terracciano et al., 2005), and test the generalizability of our theories (e.g., Benet-Martínez & John, 1998). Cultural studies, which often relay on multiple languages and samples, also offer researchers a way of dealing with classic methodological issues regarding construct validity and generalizability (e.g., need to control for possible confounding variables such as SES or language proficiency; use of multi-sample, multi-trait, multi-method designs, etc.).

Early cultural research in psychology was problematic in that it was often based on three implicit or explicit faulty assumptions (Padilla & Lindholm, 1995): (1) the use of White college samples as the standard against which other groups should be compared; (2) assuming that psychological instruments and theories are universally applicable across cultural groups or can

be used with small adjustments; and (3) neglect of possible confounding variables such as social class, education, gender, or English proficiency. This chapter challenges these assumptions and offers suggestions for conducting valid cross-cultural personality research.

II. Theoretical Issues

Most cultural studies in personality are concerned with one or both of the following (interrelated) two questions: As people of varying cultures and ethnicities, how are we different and how are we alike? How do culture and ethnicity shape our identities and personalities? Notice that the first question deals with the issue of *differences/universality* in personality, while the second question is concerned with the origin of and *processes* behind these differences. Regardless of the questions at hand, any personality psychologists doing cultural work should be familiar with the following conceptual and definitional issues: (1) Differences between culture ethnicity, race, and social class; (2) Independence between culture and personality; (3) The Emic vs. Etic debate; and (4) Differences between cultural and cross-cultural approaches.

Awareness of above issues can help cultural personality researchers frame their questions properly. Further, these issues inform the methodological considerations reviewed in sections III to VII of this chapter.

Culture, Ethnicity, and Race

Culture, ethnicity, and race are three different constructs, yet researchers often used them interchangeably. Confusion or oversights with regard to these categories results in research findings that are difficult to compare and inappropriate hypotheses and discussions. *Culture* is the broadest construct of the three and was defined earlier in this chapter. *Ethnicity* is a central component of culture. Ethnicity is neither simple nor clear-cut but it entails one or more of the following: common background or social origins, shared culture and traditions that are distinctive, maintained between generations, and result in a sense of identity and group membership, and shared language or religious tradition (Senior & Bhopal, 1994). *Race* refers to

shared genetic heritage, expressed by common external physical characteristics such as facial features, skin color, and hair texture. Because the construct of race was developed as a social classification system into which certain populations were categorized, race as an explanatory variable (at least in psychology) has been and continues to be controversial and dangerous. Accordingly, most social scientists use the more inclusive concept of ethnicity, which encompasses elements of race (for some groups only) and culture (Phinney, 1996).

A common mistake is to use the terms ethnicity and race interchangeably; but as said earlier, ethnicity usually implies 'shared identity and cultural ancestry' and race does not. Take the ethnic label 'Hispanic' or 'Latino,' for instance. Racially speaking, Hispanic individuals can be White (e.g., Spaniards or Argentineans), Black (e.g., individuals from Cuba or the Dominican Republic), or Native-American (e.g., many Mexicans and Guatemalans). Still, Hispanics define a univocal (yet admittedly very broad) distinct ethnic group because of their shared linguistic, religious, and historical traditions (i.e., predominance of the Spanish language and the Catholic religion, and being colonized by Spain).

Because of their shared elements, ethnicity and culture are also confused. Culture encompasses macro-level processes and deals specifically with the values and norms that govern and organize a group of people (e.g., capitalistic culture), defining characteristics and behaviors that are deemed appropriate or inappropriate for an organized group (e.g., American business customs). Culture also specifies the context and environment (i.e., a specific place, time, and stimuli) in which ethnicity exists. Obviously, not all individuals sharing a common 'cultural space' (e.g., US) have the same ethnicity (e.g., Hispanic, Asian-American, African-American).

The above definitional issues translate into some specific methodological recommendations. First, ironically, culture is rarely measured in most cultural studies (i.e., nationality is instead used as a proxy for culture). However, the inclusion of measures of cultural identification and/or culture-specific values and behaviors provides researchers with a tool to

conduct possible mediational analyses of their effects (or lack of thereof). Second, when dealing with ethnicity as an explanatory variable, it is important to measure each its three components (Phinney, 1996): (1) cultural values, attitudes, and behaviors associated with it; (2) strength of ethnic identification or group membership; and (3) minority status experience (e.g., discrimination, prejudice). Finally, given the growing numbers of individuals who are multi-racial and/or multicultural, researchers should be careful not exclude these individuals from participation.

Mutual Constitution of Culture and Personality

Although many studies have established that cultural forces influence the expression of personality (i.e., culture→personality effects), almost no attention has been given to the processes by which personality may in turn influence culture (personality→culture effects). Evidence from recent studies shows that our personalities shape the cultural contexts in which we live by influencing both micro- (e.g., personal spaces, music preferences, content and style of personal web pages, etc.; Gosling et al., 2002; Rentfrow & Gosling, 2003; Vazire & Gosling, 2004) and macro- (e.g., political orientation, social activism, etc.; Cole & Stewart, 1996; Jost et al., 2003) cultural elements. Thus, future cultural work in personality may benefit from using designs where researchers also explore personality effects on culture.

A related limitation of many cultural personality studies is their explicit or implicit conceptualization of culture and personality influences as unrelated forces that shape people's lives in a largely independent fashion. Even in those few studies where personality and cultural variables have been considered simultaneously (e.g., Kwan, Bond, & Singelis, 1997; Schimmack et al., 2002), the possible links between these two kinds of variables is not explicitly acknowledged. One factor contributing to this tendency to treat culture and personality as separate effects is researchers' preference for designs where culture is operationalized exclusively as an 'objective,' exogenous variable (e.g., country of birth, race, ethnicity) that

influences (i.e., moderates) the phenomenon of interest (e.g., link between personality and well-being). However, this traditional separation of cultural and personality influences is at odds with recent cultural psychology views that emphasize the inseparability and mutual constitution of psyche and culture (Aaker, Benet-Martínez, & Garolera, 2001; Markus & Kitayama, 1998; see also Church, 2000; for a review).

One example of a study examining the joint influence of culture on personality is Benet-Martinez and Karakitapoglu's (2002) cultural study on well-being. This study explored the following two questions: Do cultural syndromes, such as individualism and collectivism (Triandis, 1996), predict variations on broad personality dispositions, which in turn predict well-being? (i.e., personality as a mediator of the relationship between culture and well-being). Or rather, do personality traits drive the internalization of individualism and collectivism, which in turn relate to different levels of well-being? (i.e., cultural values as mediators of the relationship between personality traits and well-being).[2] Results from this study supported the first model (personality as a mediator between cultural values and well-being).

Emic vs. Etic Debate

A key notion in cultural research is the distinction between *imposed-etic* (imported) and *emic* (indigenous) and approaches to data collection (Berry, 1980). The imposed-etic approach (the most commonly used until recently), involves the use of instruments that are either imported in their original form or translated into the local language. This approach is economical and appropriate when the researcher's only goal is to examine how the psychometric properties of a particular 'measure' (notice that I do not say 'theory' or 'construct') generalize to other cultures. However, the imposed-etic approach has a key limitation: it precludes researchers from making valid conclusions about the cross-cultural status of the construct of interest; that is, its definition, nomological network, and prevalence in the other culture. The main problem with the imposed-etic approach is that it assumes that the construct under study (e.g., personality structure) is

defined in the culture of interest culture in exactly the same way as it is in the culture where the construct and measure were developed (e.g., in terms of five dimensions similar to those captured by the FFM); thus, the culture-specific elements of the construct (i.e., its *emic* meaning) are likely to be lost when using translated instruments.

Emic approaches, on the other hand, explore a particular psychological construct from within the cultural system. With the emic approach, instruments, and theories indigenous to the target culture are developed by relying on a systematic process (through focus groups, interviews, content analyses of popular media, or culturally-informed traditional scale development methods) that generates a set of indigenous attributes and stimuli.

The advantages of the imposed-etic strategy is that is cheap and it makes cross-cultural comparisons statistically feasible (given that quantitative judgments of similarity require stimuli that are equivalent); however, as said earlier, the imposed-etic strategy may distort the meaning of constructs in some cultures or overlook their culture-specific (emic) aspects, and thus results from cross- cultural studies that rely on this method are not fully interpretable. On the other hand, an emic-strategy, where the construct identified and measured from scratch, is well-suited to identify culture-specific aspects of a construct (i.e., it is ecologically valid), but it is expensive and renders empirical comparisons across cultures very difficult (see Church, 2001 for an excellent review of all the issues to consider in deciding between imposed-etic and emic measures).

One solution to the emic-etic debate has been to pool both approaches into what is known as a *combined emic-etic* approach (see Benet-Martinez, 1999; Yang & Bond, 1990; for illustrations of this approach). The application of a combined emic-etic approach involves the following steps: (a) identifying the emic (indigenous) elements of the construct (again, through focus groups, interviews, etc) in the target culture, and developing and administering measures that adequately tap these constructs; (b) administering translated measures of the construct (i.e.,

imposed-etic tests) in addition to the emic measures; and (c) statistically assessing the specificity and overlap between the imported and indigenous measures. This last step is key in that it allows researchers to quantify how well imported and indigenous constructs overlap/differ, and to clarify the meaning of the non-overlapping indigenous elements (e.g., see Table 2 and Figure 1 in Benet-Martinez, 1999).

Jahoda (1995) has argued that both emic and imposed-etic strategies are needed for the advancement of knowledge. In the same way that "a person is in some ways similar to all other people, in some ways to some other people, and in some ways to no other people" (Kluckhohn & Murray, 1948), some elements of social behaviors are universal (e.g., ways in which societies regulate group sharing and power; Fiske, 1991), others are common to a group or type of cultures (e.g., such as Western cultures' emphasis on the self), and some aspects are unique to a culture (e.g., Latin trait *simpatía*).

Cultural vs. Cross-Cultural Psychology

Researchers are often confused with the distinction between cross-cultural and cultural psychology. It may seem as if the main difference between these two approaches is that the first involves cultural comparisons while the second does not; however, the distinction is not so simple. These two approaches have relatively distinct conceptual, methodological, and historical elements (Greenfield, 2000), although at times the differences between these two camps have been overemphasized. Undoubtedly, both approaches share an overarching concern with the understanding of how cultural factors affect behavior.

Church's (2001) recent review of these two traditions clarifies their differences while providing ideas for possible synergy. He notes that cross-cultural studies typically have the following features: (a) a focus on individual differences (particularly personality traits); (b) comparisons of multiple cultures in the search for cultural universals, or culture-specifics along with universals; (c) conceptualization of culture as a variable "outside" the individual

(e.g., ecology, economic structure, value system) that influences personality and behavior; (d) use of traditional, standardized (i.e., context-free), psychometric scales and questionnaires; and (e) concern with the cross-cultural equivalence of constructs and measures. The majority of studies examining cultural influences on personality fit within the cross-cultural perspective. Many of these studies share an explicit or implicit optimism regarding the universality of personality dispositions and processes, particularly the Big Five (e.g., McCrae & Costa, 1997).

According to Church (2001), studies within cultural psychology, on the other hand, are often characterized by: (a) a concern with psychological processes (vs. individual differences); (b) a focus on highly-contextual descriptions of psychological phenomenon in one or more cultures, with little expectations of finding cultural universals; (c) a conceptualization of culture and psychological functioning as mutually constitutive; and (d) an emphasis on experimental methodology, coupled with qualitative or interpretive approaches. Most studies examining cultural influences on self-processes (e.g., self-enhancement, self-concept) and social behavior (e.g., attribution, dissonance, etc.) fit within the cultural perspective. Cultural psychology also speaks to the socially constructed nature of the construct of personality (e.g., how the notions of traits and personality consistency are particularly meaningful in the West).

I believe that the boundaries between the above two disciplines will become less significant as the old debates between social and personality psychology about the meaning and status of the construct of personality finally die out, as new generations of culturally-savvy psychological researchers are trained, and as the processes by which culture influences behavior become more understood. In fact, many studies are starting to combine features from both approaches, focusing on individual differences while supporting a view of culture and personality as mutually constituted (e.g., Aaker et al., 2000; Luo & Gilmour, 2004; Oishi & Diener, 2001). Implicit in these recent personality studies is the view that personality variables

(e.g., self- and other-ascribed traits, self-concept, well-being, goals) are often inseparable from cultural processes in that the ways that situations are framed and experienced and the factors a person brings to a situation (e.g., expectations, values, etc.) are cultural products themselves.

Lastly, in my view, the future of both cultural and cross-cultural psychology does not only rest as in these camps' integration but also in their ability to respond to the theoretical and methodological challenges posed by the growing phenomenon of multiculturalism. Both cultural and cross-cultural psychologists often assume that culture is a stable, uniform influence, and that nations and individuals are culturally homogeneous. But rapid globalization, continued massive migration, and the resulting demographic changes have resulted in social spaces (schools, homes, work settings) that are diverse culturally-speaking, and in growing number of individuals who identify with, and live in more than one culture (Hermans & Kempen, 1998; Hong et al., 2000). Current and future cultural studies need to develop theoretical models and methodologies that are sensitive to the multiplicity and malleability of cultural meaning within and between individuals.

III. Types of Research Questions

A personality researcher interested in how culture influences self-esteem may ask him/herself the following questions: Is self-esteem a meaningful psychological variable in non-Western societies, and if so, does it mean the same? Do individuals in Western cultures report higher levels of self-esteem than individuals in so called collectivist societies, and if so, why? Do traditional measures of self-esteem replicate well in other cultures? Is the link between self-esteem and happiness universal? Do adolescent girls in every society suffer from lower levels of self-esteem compared to boys, and is the drop in girls' self-esteem around adolescence universal? Notice that five these questions, although interrelated, deal with different issues (i.e., they ask questions about cultural differences in meaning, prevalence, nomological network, and processes in self-esteem, respectively) and thus call for different methodologies.

According to Van de Vijver and Leung (2001), most types of studies and designs in cultural research can be described in terms of three relatively independent dimensions or concerns: (1) level-oriented vs. structure-oriented studies (the most basic dimension), and (2) hypothesis-driven vs. exploratory studies; and (3) contextualized vs. non-contextualized studies (as Van de Vijver and Leung note, this taxonomy describes prototypes, and most cultural studies fit into more than one category). Level-oriented studies are mainly concerned with questions of universality vs. difference with regard to a certain personality dimension. Early cross-cultural research relied heavily on this approach; these studies found robust cultural differences in key personality variables such as locus of control (Smith, Trompenaars, & Dugan, 1995), self-concept (Bond & Cheung, 1983), or personality traits (Eysenck & Long, 1986). In structure-oriented studies, on the other hand, the emphasis is on examining cultural differences with regard to associations among variables (i.e., examining the cultural invariance of correlational patterns, regression functions, factor structures, and causal relationships). Studies using this approach have examined issues such as whether locus of control has the same correlates in different cultures (Hui & Triandis, 1983), whether the Five Factor Model is generalizable to all cultures (McCrae et al., 1998), whether self- and relational-esteem are related to life satisfaction in the same way in Hong Kong and in the USA (Kwan et al., 1997), and whether experimental manipulations of success vs. criticism produce similar increases/decreases in self-esteem in Japan and the US (Kitayama et al., 1997).

Cultural studies can also be organized along two additional dimensions: whether the purpose of the study is mainly hypothesis-testing or exploratory, and whether or not contextual factors are measured. Contextual factors are variables that can be used to validate the interpretation of cross-cultural differences, and can be either individual-related (e.g., gender, generation status, or psychological characteristics such as values or personality) or, more commonly, culture-related (e.g., GNP, type of religious or economic structure, or country-level

scores on individualism-collectivism). Variability in the above two dimensions (exploration vs. hypothesis-testing and inclusion of contextual variables) leads to four different types of studies: psychological-difference, generalizability, ecological linkage, and contextual-theory studies (see Table 1 in Van de Vijver & Leung, 2001; note that both structure- and level-oriented studies are possible in each of these four types of research).

Psychological-difference studies are mainly exploratory and do not include contextual variables. Many early comparative personality studies fit within this category (e.g., Benet & Waller, 1995; Eysenck & Long, 1986; McCrae & Costa, 1997). Generalizability studies specify and test hypotheses concerning cultural differences/similarity in the absence of contextual variables (e.g., Tafarodi & Swann, 1996). As pointed by Van de Vijver & Leung (2001), a commendable feature of most psychological difference and generalizability studies is their psychometric rigor, but they are also limited because of their neglect of contextual variables, and over-reliance on imposed-etic designs.

Ecological linkage studies explore possible explanatory variables for cultural similarities and differences reported in the literature. Such studies are usually conducted at the culture level (i.e., work with aggregate country scores) and include a large number of contextual variables. Examples of ecological-linkage personality studies are McCrae's (2001) study seeking associations between aggregate FFM personality traits and cultural variables (e.g., GNP, country scores on Individualism and Power Distance values); and Van Hemert's et al., (2002) study exploring the links between depression and subjective well-being and various relevant country characteristics (e.g., objective and relative living conditions, beliefs and values concerning happiness). Finally, studies with the most theory-building potential, according to Van de Vijver & Leung (2001), are contextual-theory studies, which are hypothesis-driven and include contextual variables. For instance, Tsai and Levenson's (1997) study tested several hypotheses regarding possible cultural differences in emotional responding during conflict between Chinese

Americans and European Americans, while also examining the role of acculturation status in explaining these differences.

Van de Vijver and Leung's (2001) taxonomy of cultural studies provides a useful and long-needed conceptual organization of traditional comparative cross-cultural and cultural studies (i.e., studies comparing two or more cultural groups). However, this taxonomy is limited in that it does not include process-oriented studies, that is, studies concerned with cultural processes and dynamics within single groups or within individuals (e.g., acculturation and biculturalism studies). And yet, process-oriented studies, particularly those relying on longitudinal designs, are instrumental in elucidating the dynamic interaction of cultural and personality factors in individuals' lives. For instance, recent process-oriented studies have examined the influence that personality traits have in the adjustment and identity structure of acculturating individuals (e.g., Benet-Martínez & Haritatos, 2005), and the personality changes that result from immigration (e.g., McRae et al., 1998).

IV. Sampling Issues

Sampling of Cultures

Very commonly, psychologists rely on samples of *convenience* for their cultural studies. This choice is typically driven by cost, logistical, and collaborator availability considerations (Church & Ortiz, 2005). In studies using samples of convenience, the choice of culture is not theory-driven and the questions and conclusions are often haphazard. Samples of convenience are often found in psychological differences studies. Unfortunately, this practice has lead to an over-representation in mainstream cultural research of modernized societies (e.g., US, Canada, Japan, Korea) and cultures that are relatively similar to each other (i.e., cultures where researchers live or access easily).

A *systematic* selection of cultures is the most optimal approach. Selection of cultures is purposeful (i.e., theory-driven) in that the cultures vary on the construct of primary interest (e.g.,

high vs. low emphasis on individual choice) or construct expected to mediate the cultural difference of interest (e.g., interdependence vs. independence), while attempting to control for extraneous variables (e.g., literacy levels, GNP, religion, climate) (Church & Ortiz, 2005).[3] But when selecting cultures systematically, how many cultures should be included? Because observed differences between cultural groups can rarely be explained in terms of a single explanatory construct, Aaker et al. (2001) and Norenzayan and Heine (2005) propose a triangulation method where at least three cultural groups, each representing variations on two possible explanatory constructs, are used. For instance, in their study of cultural differences on 'brand personality' (the abstract and instrumental qualities ascribed to a commercial brand) Aaker et al. (2001) selected the US, Japan, and Spain; the rationale behind this selection was two fold: (1) these three cultural groups vary on the socio-cultural dimensions of individualism (US vs. Spain and Japan) and affective autonomy/expressiveness (Spain and US vs. Japan), two value orientations relevant to the perception of commercial brands; (2) at the same time, US, Japan, and Spain are similar in their approach to, and resources devoted for advertising (a possible confound).[4]

Finally, some cautionary words: researchers should be careful with the common (explicit or implicit) practice of setting Western cultural groups as control group or standard of comparison against which the other cultural groups' results will be interpreted. The problem does not lay in using Western samples --which is both useful and understandable given that most cultural research is conducted in Western universities-- but in the interpretation of differences. Most of the world's population lives in collectivistic rather than individualistic societies, and thus behaviors and values such self-effacement, filial piety, or inter-dependence do in fact represent the norm. Further, discussions of difference should be wary of language implying the 'othering' or 'exoticization' of non-Western individuals.

Sampling of Participants

Within-culture diversity with respect to ethnicity, language, religion, and social class is very common. Unfortunately, however, cultural researchers often simply assume that the study participants are univocally representative of the cultures of interest, and then, when differences are found, researchers interpret these differences as "cultural." There is often a trade-off between representativeness of samples *within* cultures and equivalence of samples *between* cultures (Church & Ortiz, 2005). For instance, researchers often rely on college students to facilitate between-culture comparisons, despite the fact that college students in most cultures are more similar than they are different with regard to preferences (e.g., value for self-expression and independence) and level of affluence. A systematic approach to sampling of participants can be used to tease apart these issues. For example, by including both high and low SES subgroups within each culture, researchers will be able to interpret between-cultural group differences as cultural--vs. class or education differences (Church & Ortiz, 2005). Using this approach, Haidt and colleagues (Haidt, Koller, & Dias, 1993) found larger differences in morality reasoning between different SES groups than between cultures (Brazil and U.S.). Systematic participant sampling can also be used to rule out other powerful confounds such as linguistic ability, generation status, immigrant vs. colonized groups, acculturation level, etc. However, as Matsumoto (2000) points out, 'non-cultural' variables (e.g., demographic characteristics) are often intricately blended with culture and cannot be eliminated or controlled in the study. In these cases, researchers should conduct statistical tests to examine the contribution of these characteristics to their variables of interest, and temper their interpretations accordingly.

V. Choice and Adaptation of Instruments

Item Generation and Translation Issues

As introduced earlier, when faced with the task of measuring a particular psychological construct in another culture, researchers can use one of three strategies: (1) rely on translations of already existing measures of the construct of interest (imposed-etic approach); (2) use locally-derived instruments capturing the indigenous conceptualization of the construct (emic approach); or (3) use a combination of both types of measures (combined emic-etic approach). The emic approach is of course more labor intensive and involves a first step where the local meaning, expressions, and behavioral correlates attached to the construct of interest are identified, and a second step where the researcher and his/her team write items that adequately measure these issues. Some techniques successfully used for the first step are: (a) examining how the local media, and the popular and academic literature define the construct; (b) polling local psychologists and experts; (c) conducting focus-group discussions and in-depth interviews with relevant samples; and (d) sampling, partially or completely, the local language (in the case of taxonomic personality studies). Ideally, two or more of these strategies should be used in combination (e.g., Kim, Atkinson, & Yang, 1999). Further, in the item-writing phase, researchers should place particular effort in writing statements that are clear and could potentially be translated into other languages. The later is particularly important if the researcher plans to later compare the predictive validity of two different measures tapping different culture-specific definitions of a construct (e.g., American and Chinese notions of agency) by administrating both instruments to participants in both cultures. Brislin (1986) developed a set of guidelines for writing new items and modifying existing ones. These guidelines are summarized in Table 1.

When translations of already existing instruments are necessary, researchers should use the translation-back-translation procedure (Brislin, 1986). In this iterative method (see Figure 1), one or two bilingual individuals (ideally experts on the construct of interest) undertake the translation of the instrument from the original language into the new language. Using the same

dictionaries, a second set of bilingual experts independently translate these materials back into the original language. The combination of (a) examining the back translated versions, (b) discussions between translators, and (c) back-and-forth translations should lead to a final set of translated items that are symmetrically translatable to the original language counterparts.

After the translation-back-translation procedures, the researcher may wish to conduct a small bilingual study to pilot test the translation accuracy of the new instrument. In this bilingual design, bilingual participants complete both language versions of a questionnaire simultaneously (see Benet-Martínez & John, 1998; for a detailed illustration of this method). Dissimilar item statistics between the two versions are usually indicative of poor translations (although, see Ramirez-Esparza et al., in press). Bilingual designs have important advantages over monolingual designs because they can help unconfound the effects of language and sample differences. Notice that the bilingual design is an extension of the multitrait-multimethod approach (Campbell & Fiske, 1959), because different kinds of 'method' (i.e., language) and 'trait' (i.e., construct) effects can be tested (see Figure 1 in Benet-Martinez & John, 1998). Most recently, a set of psychometric techniques known as Item Response Theory (IRT) has become a popular and effective tool for examining cross-linguistic equivalence (e.g., Ellis, 1989).

As the above discussion attests, the cultural and linguistic challenges involved in the choice and adaptation of instruments for cultural research are not trivial; these challenges, however, can be greatly alleviated by including members of the community or culture of interest in the research team. This is strategy seems to be particularly advantageous when these individuals are included as true research partners, not just as translators or interviewers (Matsumoto, 2000).

Measurement and Conceptual Equivalence

Perhaps the most crucial methodological issue in cultural research is demonstrating the equivalence of the conceptual meaning and methodological operationalization of variables across

the cultures under study. This equivalence is an indispensable requirement for valid cross-cultural comparisons. Cultural differences in either the conceptual definitions of variables or their measurement are very common; and researchers need to be aware of these biases' origins, consequences, and available methods to control them (see Van de Vijver & Leung, 1997, for an excellent review of these issues). The notions of 'equivalence' and 'bias' are of course interrelated; equivalence can be seen as the absence of conceptual and measurement biases.

Van de Vijver and Leung (1997) describe three kinds of bias in cross-cultural research: construct bias, method bias, and item bias. Possible sources for each of these types of biases are summarized in Table 2. Because these biases are discussed in great detail in Van de Vijver and Leung's book, I will only provide short definitions and some examples for them. *Construct bias* exists when the definitions or behavioral markers of the construct being measured do not fully overlap across cultures. For instance, the concept of 'good student' is likely to be defined very differently across cultures because notions such as what constitutes good grades, sustained academic effort, appropriate amount of daily homework, and good classroom behavior vary widely across nations (Stevenson et al., 1990). Researchers should be particularly wary of the fact that, even when the constructs are defined similarly, construct bias may exist due to 'construct under-representation' (i.e., insufficient sampling of a behavioral domain).

Even if a construct is well-represented in the instrument, *method* biases may arise from particular features of the instrument or administration. Method bias comes in three forms: (a) sample bias (e.g., samples are not comparable in terms of SES, age, or familiarity with the construct of interest); (b) instrument bias (e.g., presence of differential response styles, such as acquiescence and extremity);[5] and (c) administration bias (e.g., miscommunications between researcher and participants, doing individual administration in one culture and group administration in the other). Last, *item bias* (also known as differential item functioning; DIF), can result from such factors as complex wording or translation nonequivalence for particular

items (e.g., the English and Spanish trait terms 'assertive' and 'assertivo' do not mean the same despite being linguistic cognates). The best way to reduce item bias is to ensure a good translation of the instrument items. Some available methods for investigating translation equivalence are bilinguals studies (see page 19) and cross-language studies of differential item functioning (e.g., Huang, Church, & Katigbak, 1997).

Relatedly, Van de Vijer and Leung (1997) also describe three levels of equivalence: construct (or structural), measurement unit, and scalar. *Construct or structural equivalence* is present when construct's definition and behavioral markers are cross-culturally equivalent. This type of equivalence is usually established by showing cross-cultural similarity in factorial structure and nomological networks by means of multi-group confirmatory factor analyses and path analyses (e.g., Benet-Martinez & Karakitapoglu-Aygun, 2002; Katigbak, Church, & Akamine, 1996). *Measurement unit equivalence* is present when the measure has the same unit of measurement across cultures (but perhaps different origins). *Scalar equivalence*, on the other hand, is present when the measure has the same measurement unit AND origin across cultures. These last two types of equivalence are the hardest to establish, and researchers often erroneously claim scalar equivalence after establishing construct equivalence. The presence of item bias makes scalar or full score comparability questionable.

In conclusion, researchers should be aware that comparisons of means across cultures without demonstration of conceptual and measurement equivalence can be misleading and lead to erroneous cultural stereotypes. Further, even after establishing equivalence, researchers can have more confidence in the cultural mean differences they find when the following two conditions are met (Church, 2001): (a) the cultural differences are replicated with different samples and procedures; and (b) the cultural differences are predicted and theory-driven. Alternatively, researchers may choose to rely on emic, culture-specific measures either if the study is exploratory and notions of universality or cross-cultural comparability are not as

important, or when the likelihood of obtaining measurement equivalence across cultures is small (Church, 2001).

VI. Other Procedural Issues

Experimenter's Characteristics

A couple of interrelated procedural issues that often emerge during the administration of instruments in a cross-cultural context also need to be considered; namely the cultural background of the experimenter and the nature of the interaction between experimenter and participant (see Van de Vijver & Leung, 1997, for a review). Specifically the mere presence of a culturally different person can strongly affect respondents' behavior by eliciting distrust, timidity, or display of certain demand characteristics. Misunderstandings between experimenter and participants can also give rise to administration problems, and as said earlier, unambiguous communication is a prerequisite for adequate instrument use across cultures. Because of these two issues -- the experimenter's cultural and linguistic background-- cultural studies often rely on local individuals as experimenters or test administrators (notice that this recommendation does not imply that cultural groups should only be studied only by researchers who belong to the same cultural/ethnic group).

Ethical and Political Concerns

The ethical and political issues involved in the treatment and counseling of cultural, ethnic, and linguistic minority populations have been discussed quite extensively (e.g., APA's 1993 "Guidelines for Providers of Psychological Services to Ethnic, Linguistic, and Culturally Diverse Populations"); yet the role of these issues in cultural research has received much less attention, perhaps because non-US and non-White cultural groups have only recently begun to be studies by mainstream psychology. For instance, many new cultural researchers may not be aware that most American research institutions stipulate that research conducted in foreign countries remains under the institution's purview and guidelines. Typically, cross-cultural

research projects must have been approved by the foreign equivalent of an IRB before they can be approved by the investigator's local IRB. Where there is no equivalent board or group, investigators must rely on local experts or community leaders to provide approval. These principles allow for the fact that, while the investigator's institution cannot impose its standards for written documentation on the cultures under study, the standards for ethical conduct of research (e.g., consent process) should be applied.

Further, ethical considerations in doing cultural research call for a careful consideration and respect of the social (e.g., SES), cultural (e.g., intra-group differences in language and values), and political (e.g., immigrant status) contexts of the cultural groups under study. Overlooking the above issues not only may be unethical but may compromise the quality of the data and its interpretation. For instance, participants who are unfamiliar with research practices and/or have different cultural expectations about what these entail, may feel reluctant to disclose information about risky or highly personal behaviors out of decorum or concern that such disclosure may harm their or their family's reputation and legal status. Further, because of these possible differences in expectations and beliefs, some participants may mistakenly believe that the researcher can actively assist them in obtaining help for personal and communal issues revealed during the course the project (Fisher et al., 2002).

As it was the case when dealing with potential linguistic bias (see page 19), an important resource for dealing with these ethical and political issues is to include the opinions of individuals with knowledge of and experience with the culture and behaviors that are the target of investigation. Respectful and successful community and participant consultation often depends on establishing a relationships of trust early in the research design phase, and to rely on this expertise continuously through the data interpretation, implementation, and dissemination phases (Fisher et al., 2002).

Priming Effects

Cultural researchers need also be aware of potential priming effects in their procedures. Specifically, the order and wording of certain questions and stimuli may move participants' responses in a direction that facilitates or difficults the hypothesized behavior or differences (see also Footnote 5). For instance, merely asking people to think for two minutes about what they have in common with their family and friends, often functions as a prime that shifts people toward collectivism, whereas asking them to think of what makes one different from family and friends shifts them towards individualism (Trafimow, Triandis, & Gotto, 1991). This cultural 'frame-switching" is even more pronounced among bicultural individuals (Hong et al. 2000).

Do the above findings mean that many reported cultural differences are superficial or not meaningful? No. Recent cultural studies suggest that, at least at the level of basic social values and schemas (or what Bond calls 'social axioms;' Bond et al., 2004), all individuals regardless of their cultural background possess both 'individualist' and 'collectivist' cognitive structures (i.e., schemas about agency and uniqueness, and schemas about interrelatedness and obligation), and yet the chronic accessibility to these structures differs greatly cross-culturally because of national differences in institutions, discourse and practices (e.g., prevalence of the protestant ethic, value of free agency and competition, capitalism; Kitayama et al., 1997). Still, as said earlier, researchers should be aware that the presence of certain stimuli (e.g., images or words associated with a particular cultural worldview) may make certain value schemas or mindsets temporarily accessible and thus influence participants' responses.

Reference-Group Issue

Several cultural researchers (Heine et al., 2002; Peng et al., 1997) have argued that some of the null results obtained from cross-cultural comparisons of means from attitude, trait, and value measures (e.g., Oyserman, Coon, & Kemmelmeir, 2001) are problematic and non-interpretable because of cultural differences in participants' choice of a comparison standard or what is known as the 'reference-group' effect; that is, those in one culture may compare

themselves with different others and standards than do those in another culture, thus potentially confounding cross-cultural comparisons.

The above researchers argue that reliable and coherent cultural group differences are more likely to be observed when cultural differences are made salient by placing contrastive cultural paradigms in juxtaposition, such as when Japanese and Canadians are asked to use each other as a reference group to calibrate their self-ratings (Heine et al. 2002), when individualism scores and collectivism scores are pitted against each other (Schimmack et al. 2002), or when using scenarios understood by people from both target cultures (Peng et al., 1997). Oishi et al. (2004) provide an excellent conceptual and empirical analysis of this issue and the advantages and disadvantages of each of these aforementioned methods.

VII. Data Analyses and Interpretation of Results

Data analysis in cultural research, like in any other types of research, involves a strategic choice of statistical techniques made on the basis of substantive considerations such as the research questions or hypotheses, sample size, data type etc. Analysis and interpretation of cultural data, however, often involves the additional task of demonstrating measurement and conceptual equivalence (see Section 5 of this chapter).

Preliminary Analyses: Cultural Response Sets and Item biases

A careful examination of the psychometric characteristics of instruments is an important first step in the analysis of cultural data. A preliminary (but incomplete) method of examining possible biases is to compare the instruments' reliabilities across the cultures under study (Van de Vijver & Leung, 1997). If significant differences are observed, which is common, their source should be explored (e.g., examine item-total correlations for each scale and sample). Of course such differences can be produced by item bias due to bad translation, administration problems (e.g., experimenter effects or low inter-judge reliability), sample characteristics (e.g.,

cultural differences in test familiarity or education levels), and differential response styles (e.g., acquiescence or social desirability).

Cultural response sets are tendencies for members of a given culture to use certain parts of a scale when responding (e.g., social desirability biases, extreme responding, acquiescence). For instance, individuals from collectivistic cultures may be reluctant to use the extreme end points of a scale, consistent with a cultural resistance to "stick out." Alternatively, individuals less familiar with likert scales may be more inclined to use the endpoints to signal yes/no or true/false responses. Overall, work on this topic seems to suggest that persons who are older, less educated, or come from lower socioeconomic strata are more likely to display response styles (Van de Vijver & Leung, 1997). Cultural response sets cloud the interpretation of results because any differences found among cultures may reflect these response tendencies rather than actual cultural differences on the items and constructs of interest. However, correction for the biases does not always reduce these differences (Grimm & Church, 1999).

Prior to any statistical analysis, the researcher has also to decide if the data should be standardized within each cultural group, and if so, which standardization procedure is to be used (Leung & Bond, 1989). Generally speaking, standardization is defined as the computation of z scores ($z = [X - MI /S$, in which X is the score to be standardized, M is the mean of the cultural group, and S is its standard deviation). Typically researchers use standardization to reduce or eliminate unwanted cultural differences due to response sets. However, when scores are standardized per cultural group, all (true and biased) cultural differences in means, standard deviations, or both are eliminated. Thus, it is appropriate to use this method prior to factor analyses (e.g., as when the goal is to compare factor structures; Yang & Bond, 1990; Benet-Martinez & Waller, 1997) but not for cross-cultural mean score comparisons.

When the demonstration of lack of item bias is important, differential item functioning statistical techniques such as analyses of variance, IRT, or Mantel-Haenszel Statistic (see Van de

Vijer & Leung, 1997; for a review of these procedures) provide a much more rigorous tests than classical item-statistics. Ramirez-Esparza et al. (2005) for instance, used these kinds of item-bias analyses to rule out translation anomalies before interpreting differences in responses to Big Five questionnaires in English and Spanish within bilinguals. But when item biases are detected, how should the researcher deal with them? Item biases can be seen as an indication that an instrument or particularly item is not inadequate for cross-cultural comparison. Such approach is prudent but sometimes too restrictive because item biases are likely in any study of highly dissimilar cultural groups (Van de Vijver & Leung, 1997). Alternatively, item bias can be seen as providing clues to possible meaningful cultural idiosyncrasies, although the appropriateness of this approach is contingent on successfully finding reasons for the presence and absence of bias, possibly aided by one or more local experts.

It is important to note that the use of item bias statistics is not free of limitations (Van de Vijver & Leung, 1997). First, it is often very difficult to identify reasons for why an item is biased. Second, different procedures for identifying bias often do not yield the same results. Third, the stability of item bias statistics is often poor. Still, value of item bias analyses should not be underestimated.

Structure- and Level-Oriented Studies

In structure-oriented studies, where the main goal is examine construct equivalence across cultures, a relatively large set of statistical techniques is available. The most frequently used is exploratory factor analysis followed by target rotations and the computation of an index of factorial agreement across cultural groups (e.g., McCrae et al., 2005). One key complication in factor comparison is the "rotation problem": The spatial orientation of factors in factor analysis is arbitrary. Factor solutions obtained in different cultural groups may be rotated with regard to each other (without this rotation factor similarity will be underestimated). The problem is that one cultural group needs to be arbitrarily designated as the target factor structure, and

unfortunately most studies set the imported structure (e.g., the English NEO structure obtained with US participants) as the target towards which the other group structures will be shifted to (see discussion in page 16 of this chapter).

A more sophisticated method to test construct equivalence is to use multi-group confirmatory factor analysis (MCFA), which allows researchers to test the fit of a series of hypothesized factorial structures in two or more cultural groups simultaneously (e.g., Benet-Martinez & John, 1998). Joint confirmatory factor analysis (JCFA), in which the common underlying structure of two or more instruments is examined, is also particularly useful in studies structure-level cultural studies, particularly in those using a combined emic-etic approach (see Figure 2). Recall that a key goal of this approach is to examine the degree of overlap/difference between indigenous and imported (i.e., translated) constructs. As Figure 2 shows, JCFA techniques are optimal at revealing which (and to what extent) identified indigenous personality dimensions have cultural-specific vs. common meaning, and thus how much imported definitions of a particular construct leave out meaning that is unique to the culture under study. When the goal is to examine the cross-cultural invariance of a particular nomological network, pattern of variable associations, or causal model, multi-group path analyses are also quite advisable (e.g., Kwan et al., 1997; Benet-Martinez & Karakitapoglu, 2002). This technique (another variant of structural equation modeling techniques) allows for the examining of multiple dependent variables and direct and indirect effects.

Cross-validation is key in determining the plausibility of postulated structural and causal models within or across cultures. When a replication study is not feasible, split-sample cross-validation techniques are recommended assuming that the sample size is large enough. In this technique, the (new or hypothesized) factor structure is identified using one random half of the sample, and the second half is used to examine how well the identified structure replicates with a different sample (e.g., Benet-Martinez, 1999).

In level-oriented studies, the most frequently used statistical tests are the t test and ANOVA. As emphasized earlier, unless the presence of bias cannot be ruled out, the interpretation significant difference found through these tests may be ambiguous. Also often used in cultural research is the more complex so-called factorial designs, where in addition to cultural group, one or more independent variables such as gender, age, SES, or generations or acculturation status (i.e., key possible confounds or covariates) are included. These tests can also be achieved with multiple regression techniques, where culture is entered as a dummy-coded variable. Multiple regression analyses are particularly useful when the goal is to examine if the relative importance (i.e., beta weights) of a specific set of independent variables varies across two cultural groups.

Interpretation of Results: Individual- vs. Cultural-level

At least two levels of analysis are possible in cultural research (Leung & Bond, 1989). In the culture-level approach, culture or nation is the unit of analysis, and the results inform the characterization of cultures but not of individuals. Hofstede's (1980) classic study of cultural values across 50+ cultures was based on this approach (see these dimensions in Table 4). In the individual-level approach, the individual is the unit of analyses (see Table 3). Dimensions and results obtained at the individual level may not replicate at the cultural level and vice versa. For instance, notice that the names of the value dimensions reported in Tables 3 (individual-level) and 4 (cultural-level), as well as their personality, well-being, income correlates, are different. This indicates that both the meaning of and dynamics among these variables vary across the psychological and ecological levels. Subjective well-being, for instance, is positively associated with gross national product at the cultural level (i.e., richer countries report higher levels of well-being), but at the individual level, happiness and income do not seem to be related (Myers & Diener, 1996).

Two data analysis techniques with great potential value in cultural research and capable

of handling the above complexities are multilevel modeling (e.g., Bryk & Raudenbush, 1992; Van de Vijver & Leung, 1997) and latent class analysis (e.g., Eid & Diener, 2001). These techniques allow researchers to compare and link findings at the individual and cultural levels, and are particularly useful for identifying within-culture heterogeneity as well as between-culture differences. I believe these underused techniques have the potential of fostering a fruitful synergy between the field of personality –which has provided a wealth of information regarding individual-level psychological characteristics (e.g., traits and values)—and the fields of and anthropology or sociology, which are very informative regarding culture-level phenomena (e.g., economy, religion, and many other key demographic factors).

VIII. Conclusions

As this chapter hopefully attests, cultural research offers scientists unique benefits and opportunities (e.g., elucidate links between individual and ecological influences on personality, dispel cultural stereotypes, test theory generalizability). Cultural studies, in fact, may make us better at 'seeing' personality. Supporting this idea, anthropologist Margaret Mead observed "… the individual's inclination to respond in a certain ways is relatively stable when the cultural context is understood." (qtd. in Friedman & Schustack, 2003). In other words, by understanding the cultural backdrop of a particular behavior or script, culturally-informed personality researchers may correctly see individual differences and patterns of personality consistency and coherence where other researchers would only see situational or random variability (Oishi, 2004).

Finally, let's not forget the important societal and applied benefits of cultural personality studies. Cultural personality research offer scientists, managers, policy-makers, and the public ways to understand, manage, and benefit from the omnipresent cultural diversity that characterizes our society (Fowers & Richardson, 1996).

References

Aaker, J., Benet-Martínez, V., & Garolera, J. (2001). Consumption symbols as carriers of culture: A study of Japanese and Spanish brand personality constructs. Journal of Personality and Social Psychology, 81, 249-264.

Allport, G.W. (1954). The nature of prejudice. Cambridge, MA: Addison-Wesley.

Benet-Martinez, V. (1999). Exploring indigenous Spanish personality constructs with a combined emic-etic approach. In J. C. Lasry, J. G. Adair, & K. L. Dion, (Eds.), Latest contributions to cross-cultural psychology (pp. 151-175). Lisse, The Netherlands: Swets and Zeitlinge.

Benet-Martínez, V. & John, O. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait multimethod analyses of the Big Five in Spanish and English. Journal of Personality and Social Psychology, 75, 729-750.

Benet-Martínez, V., & Karakitapoglu-Aygun, Z. (2003). The interplay of cultural values and personality in predicting life-satisfaction: Comparing Asian- and European-Americans. Journal of Cross-Cultural Psychology, 34, 38-61.

Benet, V., & Waller, N. G. (1995). The "Big Seven" model of personality description: Evidence for its cross-cultural generality in a Spanish sample. Journal of Personality and Social Psychology, 69, 701–718.

Berry, J.W., (1980). Introduction to Methodology. In H.C. Triandis & J.W. Berry (Eds.), Handbook of Cross-Cultural Psychology: Vol. 2. Methodology. Boston: Allyn and Bacon.

Bond, M.H., et al. (2004). Culture-level dimensions of social axioms and their correlates across 41 cultures. Journal of Cross-Cultural Psychology, 35, 548-570.

Bryk, A.S. & Raudenbush, S. W.(1992). Hierarchical linear models: Applications and data analysis methods. Thousand Oaks, CA: Sage.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. Behavioral and Brain Sciences, 12, 1-49.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81–105.

Chiu, C. Y., & Chen, J. (2004). Symbols and interactions: Application of the CCC Model to culture, language, and social identity. In S. H. Ng, C. N. Candlin, & C. Y. Chiu (Eds.), Language matters: Communication, Culture, and Identity (pp. 155-182). Hong Kong: City University of Hong Kong Press.

Church, A.T. (2000). Culture and personality: Toward an integrated cultural trait psychology. Journal of Personality, 68, 651-703.

Church, A.T. (2001). Personality measurement in cross-cultural perspective. Journal of Personality, 69, 979-1006.

Church, A. T. & Ortiz, F. A. (2005). Culture and personality. In V. J. Derlaga, B. A. Winstead, & W. H. Jones (Eds.), Personality: Contemporary theory and research (3nd ed., pp. 420-456). Belmont, Ca: Wadsworth.

Cole, E. R., & Stewart, A. J. (1996). Meanings of political participation among Black and White women: Political identity and social responsibility. Journal of Personality and Social Psychology, 71, 130-140.

Diener, E., Diener, M, & Diener, C. (1995). Factors predicting the subjective well-being of nations. Journal of Personality and Social Psychology, 69, 851-864.

Diener, E., Oishi, S., Lucas, R.E. (2003). Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life. Annual Review of Psychology, 54, 403-425.

Eid, M., & Diener, E. (2001). Norms for experiencing emotions in different cultures: Inter- and intranational differences. Journal of Personality and Social Psychology, 81, 869-885.

Ellis, B.B. (1989). Differential item functioning: Implications for test translations. Journal of Applied Psychology, 74, 912-921.

Fisher, C. B., Hoagwood, K., Boyce, C., Duster, T., Frank, D. A., Grisso, T. et al. (2002). Research ethics for mental health science involving ethnic minority children and youths. American Psychologist, 57, 1024-1040.

Fiske, A. P. (1991). Structures of social life: The four elementary forms of human relations. New York, NY: Free Press.

Fowers, B. J., & Richardson, F. C. (1996). Why is multiculturalism good? American Psychologist, 51, 609-621.

Friedman, H., & Schustack, M.W. (2003). Personality: Classic Theories and Modern Research (2nd edition). Boston, MA: Allyn and Bacon.

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A Room with a cue: Judgments of personality based on offices and bedrooms. Journal of Personality and Social Psychology, 82, 379-398.

Greenfield, P. M. (2000). Three approaches to the psychology of culture: Where do they come from? Where can they go? Asian Journal of Social Psychology, 3, 223-240.

Grimm, S.D., & Church, A.T. (1999). A cross-cultural study of response biases in personality measures. Journal of Research in Personality, 33, 415-441.

Haidt, J., Koller, S., & Dias, M. (1993). Affect, culture, and morality, or is it wrong to eat your dog? Journal of Personality and Social Psychology, 65, 613-628.

Heine, S. J., Lehman, D. R., Peng, K., Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales? The reference-group effect. Journal of Personality and Social Psychology, 82, 903-918.

Hermans, H., & Kempen, H. (1998). Moving cultures: The perilous problem of cultural dichotomies in a globalizing society. American Psychologist, 53, 1111-1120.

Hofstede, G. (1980) Culture's Consequences. Beverley Hills: SAGE Publications.

Hong, Y. Y., Morris, M., Chiu, C. Y., & Benet-Martínez, V. (2000). Multicultural minds: A dynamic constructivist approach to culture and cognition. American Psychologist, 55, 709-720.

Huang, C. D., Church, A. T., & Katigbak, M. S. (1997). Identifying cultural differences in items and traits: Differential item functioning in the NEO Personality Inventory. Journal of Cross-Cultural Psychology, 28, 192-218.

Jahoda, G. (1995). In pursuit of the emic-etic distinction: Can we ever capture it? In N. R. Goldberger & J. B. Veroff (Eds.), The Culture and Psychology Reader (pp. 128-138). New York: New York University Press.

Jost, J.T., Glaser, J., Kruglanski, A.W., & Sulloway, F. (2003). Political conservatism as motivated social cognition. Psychological Bulletin, 129, 339-375.

Katigbak, M. S., Church, A. T., & Akamine, T. X. (1996). Cross-cultural generalizability of personality dimensions: Relating indigenous and imported dimensions in two cultures. Journal of Personality and Social Psychology, 70, 99-114.

Kim, B. S. K., Atkinson, D. R., & Yang, P. H. (1999). The Asian Values Scale: Development, factor analysis, validation, and reliability. Journal of Counseling Psychology, 46, 342-352.

Kwan, V. S. Y., Bond, M. H., & Singelis, T. M. (1997). Pancultural explanations for life-satisfaction: Adding relationship harmony to self-esteem. Journal of Personality and Social Psychology, 73, 1038-1051.

Kluckhohn, C., & Murray, H. E. (1948). Personality in nature, society, and culture. New York, NY: Alfred A. Knopf.

LeVine, R. A. (2001). Culture and Personality Studies, 1918–1960: Myth and History. Journal of Personality, 69, 803-818.

Leung, K., & Bond, M. H. (1989). On the empirical identification of dimensions for crosscultural comparisons. Journal of Cross-Cultural Psychology 20, 133-151.

Lou, L., & Gilmour, R. (2004). Culture and conceptions of happiness: Individual oriented and social oriented subjective wellbeing. Journal of Happiness Studies, 5, 269-291.

Markus, H. R., & Kitayama, S. (1998). The cultural psychology of personality. Journal of Cross-Cultural Psychology, 29, 63-87.

Matsumoto, D.R. (2000). Culture and psychology: People around the world. Delmar, CA: Wadsworth Thomson Learning.

McCrae, R.R. (2001). Trait psychology and culture: Exploring intercultural comparisons. Journal of Personality, 69, 819-846.

McCrae, R. R., & Costa, P. T., Jr. (1997). Personality trait structure as a human universal. American Psychologist, 52, 509-516.

McCrae, R. R., Terraciano, A., & 79 members of the Personality Profiles of Cultures Project (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. Journal of Personality and Social Psychology, 88, 547-561.

McCrae, R. R., Yik, M. S. M., Trapnell, P. D., Bond, M. H., & Paulhus, D. L. (1998). Interpreting personality profiles across cultures: Bilingual, acculturation, and peer rating studies of Chinese undergraduates. Journal of Personality and Social Psychology, 74, 1041-1055.

McClelland, D. C. (1961) The achieving society. Princeton: Van Nostrand.

Myers, D. G., & Diener, E. (1995). Who is happy? Psychological Science, 1, 10-19.

Norenzayan, A., & Heine, S.J. (2005). Psychological universals: What are they and how can we know? Psychological Bulletin, 131, 763-784.

Oishi, S. (2004). Personality in culture: A neo-Allportian view. Journal of Research in Personality, 38, 68-74.

Oishi, S., & Diener, E. (2001). Goals, culture, and subjective well-being. Personality and Social Psychology Bulletin, 27, 1674-1682.

Oishi, S., Hahn, J., Schimmack, U., Radhakrishan, P., Dzokoto, V., Ahadi, S. (2004). The measurement of values across cultures. Journal of Research in Personality. 39, 299-305.

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. Psychological Bulletin, 128, 3-72.

Padilla, A. M., & Lindholm, K. J. (1995). Quantitative educational research with ethnic minorities. In J. Banks & C. A. McGee Banks (Eds.), Handbook of research on multicultural education (pp. 97-113). New York: Macmillan.

Peng, K., Nisbett, R. E., & Wong, N. Y. C. (1997). Validity problems comparing values across cultures and possible solutions. Psychological Methods, 2, 329–344.

Phinney, J. S. (1996). When we talk about American ethnic groups, what do we mean? American Psychologist, 51, 918-927.

Ralston, D.A., Cunniff, M.K., & Gustafson, D.A. (1995). Cultural accommodation: The effect of language on the responses of bilingual Hong Kong Chinese managers. Journal of Cross-Cultural Psychology, 26, 714-727.

Ramírez-Esparza, N., Gosling, S.D., Benet-Martínez, V., Potter, J., & Pennebaker, J.W. (in press). Do bilinguals have two personalities? A special case of frame switching. Journal of Research in Personality.

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. Journal of Personality and Social Psychology, 84, 1236-1256.

Roccas, S., Sagiv, L., Schwartz, S. H., & Knafo, A. (2002). The Big Five Personality Factors and Personal Values. Personality and Social Psychology Bulletin, 28, 789-801.

Schimmack, U., Radhakrishnan, P., Oishi, S., Dzokoto, V., & Ahadi, S. (2002). Culture, personality, and subjective well-being: Integrating process models of life-satisfaction. Journal of Personality and Social Psychology, 82, 582-93.

Schwartz, S.H. (1994). Beyond individualism-collectivism: New dimensions of values. In: U. Kim, H. C. Triandis, C. Kagitcibasi, S.C. Choi and G. Yoon (Eds.) Individualism and collectivism: Theory, method and application (pp. 85-119). Newbury Park, CA: Sage.

Senior, P. A, & Bhopal, R. (1994). Ethnicity as a variable in epidemiological research. Briths Medical Journal, 309, 327-30.

Soldz, S., & Vaillant, G. E. (1999). The Big Five personality traits and the life course: A 45-year longitudinal study. Journal of Research in Personality, 33, 208-232.

Stevenson, H. W., S. Y. Lee, C. Chen, M. Lummis, J. Stigler, L. Fan, & F. Ge. (1990). Mathematics Achievement of Children in China and the United States. Child Development 61, 1053-1066.

Tafarodi, R.W., & Swann,W. B., Jr. (1996). Individualism-collectivism and global self-esteem: Evidence for a cultural trade-off. Journal of Cross-Cultural Psychology, 27, 651-672.

Terraciano et al., (2005). National character does not reflect mean personality trait levels in 49 cultures. Science, 310, 96 – 100.

Trafimow, D., Triandis, H. C., & Goto, S. G. (1991). Some tests of the distinction between the private self and the collective self. Journal of Personality and Social Psychology, 60, 649-655.

Triandis, H.C. (1996). The psychological measurement of cultural syndromes. American Psychologist, 51, 407-415.

Triandis, H. C., & Suh, E. M. (2002). Cultural influences on personality. Annual Review of Psychology, 53, 133–160.

Tsai, J. L., & Levenson, R. W. (1997). Cultural influences on emotional responding: Chinese and European American dating couples during interpersonal conflict. Journal of Cross-Cultural Psychology, 28, 600-625.

Van de Vijver, F., & Leung, K. (1997). Methods and data analysis for cross-cultural research. London, UK: Sage.

Van Hemert, D. A., van de Vijver, F. J. R., Poortinga, Y. H., & Georgas, J. (2002). Structural and functional equivalence of the Eysenck Personality Questionnaire within and between countries. Personality and Individual Differences, 33, 1229-1249.

Vazire, S., & Gosling, S. D. (2004). E-perceptions: Personality impressions based on personal websites. Journal of Personality and Social Psychology, 87, 123-132.

Yang, K., & Bond, M. H. (1990). Exploring implicit personality theories with indigenous and imported constructs: The Chinese case. Journal of Personality and Social Psychology, 58, 1087-1095.

Table 1: Guidelines for Writing New Items and Editing Existing Ones (from Van de Vijver & Leung, 1997).

1. Use short, simple sentences of fewer than 16 words.

2. Employ the active rather than the passive voice, because the former is easier to comprehend.

3. Repeat nouns instead of using pronouns, because the latter may have vague referents; in English, for example, *you* can refer to any number of persons.

4. Avoid metaphors and colloquialisms.

5. Avoid the subjunctive form, with words like *could* and *would*. Many languages express this meaning in different ways, thereby putting a burden on the translator.

6. Add sentences to provide context for key ideas. Redundancy is not harmful for communicating key aspects of the instrument.

7. Avoid verbs and prepositions telling "where" and "when" that do not have a definite meaning. How many times a week do you have to see someone in order to say that you see him "often"?

8. Avoid possessive forms where possible, because it may be difficult to determine the ownership. The ownership such as "his" in "his dog" has to be derived from the context of the sentence and languages do not have similar rules for expressing this ownership.

9. Use specific rather than general terms. Who is included in "members of your family" strongly differs across cultures; more precise terms are less likely to run into this problem.

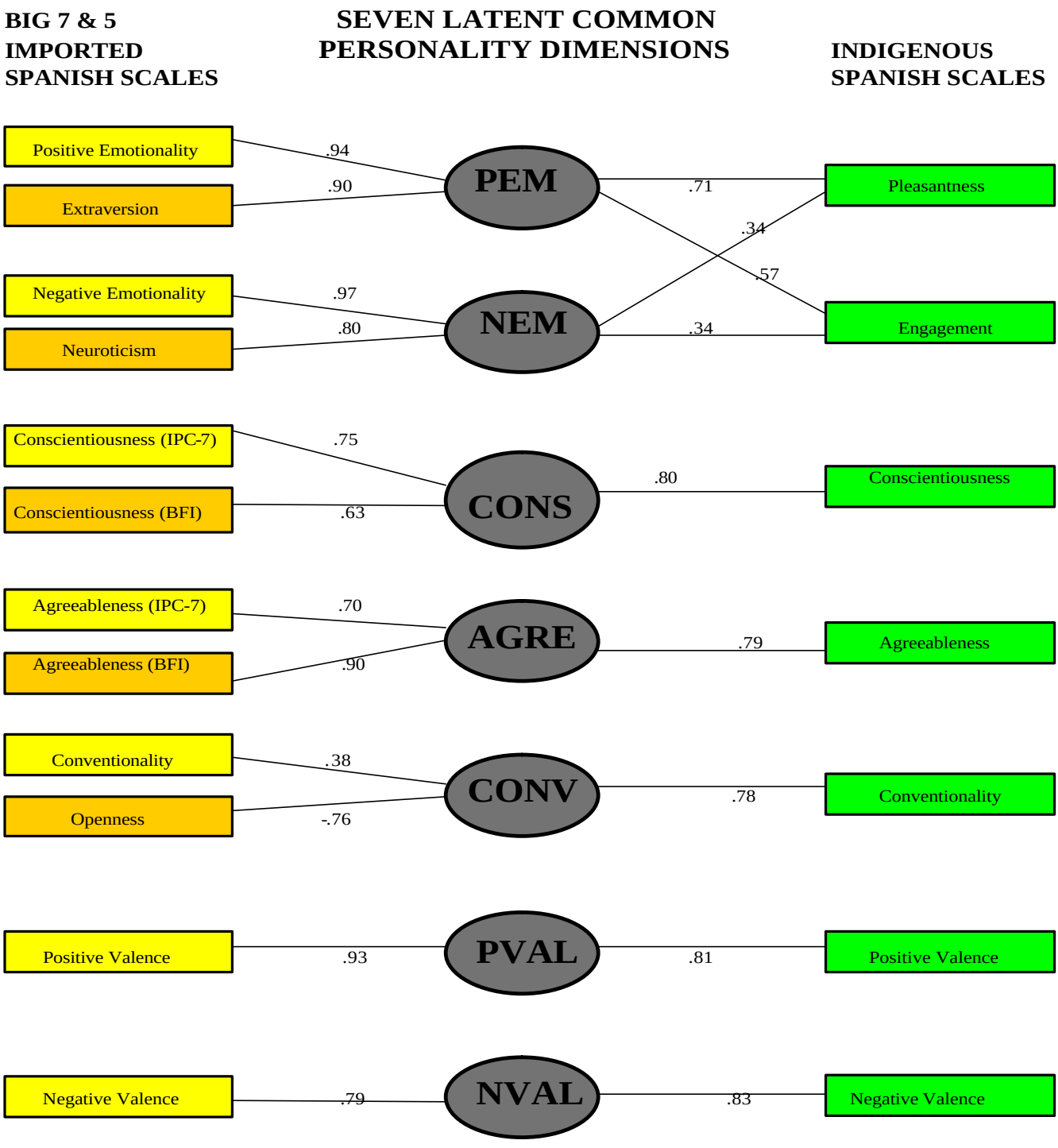
Figure 1: Translation-Back-Translation Model

Step	Description
A	Two simultaneous independent translations are made from source to target language by two bilingual experts.
B	Each target language version is blindly back-translated to the source language by two new bilingual experts.
C	The four bilingual experts involved meet with the investigators to review the back-translations, identify differences in meaning, and adapt the target language version to achieve the most accurate culturally equivalent meaning.
D	The new version is independently back-translated by two more bilingual experts.
E	A second meeting of the bilingual experts to review the new back-translations. If necessary, the process continues until the team agrees on the culturally equivalent meaning in the source- and target-language versions of the instrument.
F	Validation of the back-translated instrument by testing for reliability and equivalence using a sample of bilingual subjects.

Table 2: Overview of Types of Bias and Their Most Common Causes (from Van de Vijver & Leung, 1997).

TYPE OF BIAS	*SOURCE*
Construct	Incomplete overlap of definitions of the construct across cultures
	Differential appropriateness of (sub)test content (e.g., skills do not belong to the repertoire of one of the cultural groups)
	Poor sampling of all relevant behaviors (e.g., short instruments)
	Incomplete coverage of the construct (e.g., not all relevant domains are sampled)
Method	Differential social desirability
	Differential response styles such as extremity scoring and acquiescence
	Differential stimulus familiarity
	Lack of comparability of samples (e.g., differences in educational background, age, or gender composition)
	Differences in physical conditions of administration
	Differential familiarity with response procedures
	Tester/interviewer effects
	Communication problems between respondent and tester/interviewer in either cultural group
Item	Poor item translation
	Inadequate item formulation (e.g., complex wording)
	Item(s) may invoke additional traits or abilities
	Incidental differences in appropriateness of the item content (e.g., topic of item of educational test not in curriculum in one cultural group)

Figure 2: Example of Analyses and Results from a Combined Emic-Etic Approach in Cultural Personality Research



Note = From Benet-Martinez (1999); N = 894 Spanish college students; Seven latent personality factors (PEM = Positive Emotionality; NEM = Negative Emotionality; CONS = Conscientiousness; A = Agreeableness; CONV = Conventionality; PVAL = Positive Valence; NVAL = Negative Valence); CFI = .945 and $X^2/df = 2.3$

Table 3: Correlations Between Big Five and Cultural and Economic Variables (Individual-Level)

	N	E	O	A	C
Singelis Value Dimensions[a]					
Individualism	**-.34****	**.52****	**.52****	.11	.17
Collectivism	-.02	.01	-.23	**.33****	.09
Schwartz Value Dimensions[b]					
Benevolence	-.02	.01	-.06	**.45****	.04
Universalism	-.02	-.07	**.47****	**.15***	**-.17***
Self-Direction	-.10	.10	**.48****	**-.25****	-.01
Stimulation	-.07	**.26****	**.33****	**-.26****	**-.24****
Hedonism	-.01	**.18****	.07	**-.34****	-.05
Achievement	**-.21****	**.31****	-.06	**-.41****	**.22****
Power	.08	**.13***	**-.38****	**-.45****	.05
Security	.02	-.11	**-.29****	.06	**.22****
Conformity	.02	**-.13***	**-.34****	**.20****	**.16****
Tradition	**.12***	**-.29****	**-.29****	**.36****	-.10
Conscientiousness	-.04	**-.18****	**-.26****	-.04	**.40****
Life Satisfaction[a]	**-.47****	**.34*****	.18	**.26****	**.26****
Salary[c]	-.03	**.22*****	-.01	.01	.09

Note. [a] N = 122 US college students from Benet-Martínez & Karakitapoglu's (2002) study; [b] N = 246 Israelis from Roccas et al.'s (2002) study; SWB = Subjective well-being; [c] N = 163 US men from Soldz & Vaillant (1999) study.

Table 4: Correlations Between Big Five and Cultural and Economic Variables (Culture-Level)

	N	E	O	A	C
Hofstede Value Dimensions					
Power Distance	.20	**-.46*** **	**-.41** **	**-.31***	.11
Uncertainty Avoidance	**.30***	.07	-.03	-.02	.20
Individualism vs. Collect	.05	**.51** **	**.33***	**.37** **	-.14
Masculinity	-.14	.00	.10	.04	.03
Schwartz Value Dimensions					
Conservatism	-.20	-.02	**-.70** **	**-.51***	.15
Affective Autonomy	.13	.24	**.55** **	**.61** **	-.03
Intellectual Autonomy	.37	-.15	**.51***	**.44***	.07
Hierarchy	-.24	-.12	-.32	-.23	-.10
Mastery	-.27	-.31	.10	-.09	-.15
Egalitarian Commitment	.25	.20	**.55** **	**.44***	-.09
Harmony	.08	.09	.26	.09	.15
Life Satisfaction	.01	**.63*** **	**.35***	**.48** **	-.02
GDP	.04	**.44*** **	**.47*** **	**.46*** **	.02

Note. Adapted from McCrae & Terraciano (2005); N = 22-41 different cultures; GDP = per capita Gross Domestic Product.

Endnotes

[1] The terms “cross-cultural psychology” and “cultural psychology” refer to two different research traditions with somewhat distinct theoretical approaches, goals, and methodologies (see page X for a discussion of these issues). However, for the sake of simplicity, throughout the chapter I will use the broader term ‘cultural’ (e.g., cultural psychology, cultural research, cultural methods) to refer to both kinds of traditions and their theories and methodologies.

[2] Note that these two questions deal, if not directly, with a basic disagreement within personality psychology; namely, the opposing views that see the Big Five as largely representing either (a) endogenous and inherited basic tendencies that are independent from culture (genotypic view; McCrae & Costa, 1997), or (b) observable behavioral regularities that reflect characteristic adaptations to the socio-cultural context (phenotypic view; Saucier & Goldberg, 1996).

[3] In choosing systematically the cultures for study, researchers often rely on Hofstede's (1980, 2001) rankings of more than 50 cultures along the dimensions of Individualism-Collectivism, Power Distance, Uncertainty Avoidance, and Masculinity-Femininity. Other useful rankings of cultures along meaningful socio-cultural dimensions (e.g., values) can be found in Schwartz (1994), and Smith, Dugan, & Trompenaars (1996).

[4] Note that *random* sampling of cultures is often not feasible or desirable, but in some large-scale exploratory studies this has been an ideal approach (e.g., Buss, 1989; Diener, Diener, & Diener, 1995).

[5] Instrument biases are complex and deserve special attention, although their cultural reliability is still being debated (Grimm and Church, 1999). Also, in multilingual persons, there is evidence suggesting that the language of assessment may be a potential source of method bias either because of *cultural frame-switching* (i.e., each language primes different aspects of the

self; Ramirez-Esparza et al., in press) or *cultural accommodation* effects (i.e., the respondent answers the questionnaire in a manner that accommodates or favors the culture associated with the language being used; Yang & Bond, 1982; Ralston, Cunniff, & Gustafson, 1995).