Review of literature on the Delphi Technique

Sinéad Hanafin

2nd March 2004

Acknowledgements

Table of Contents

Introduction

The Delphi technique is a research approach used to gain consensus through a series of rounds of questionnaire surveys, usually two or three, where information and results are fed back to panel members between each round. We are proposing to adopt this technique to facilitate the development of a national set of child well-being indicators. The purpose of this paper is to provide an overview of the Delphi technique as a research methodology and to highlight key issues, advantages and disadvantages of this approach to achieving consensus. The paper also sets out key questions arising in respect of the proposed study on the development of a national set of child well-being indicators in Ireland. In doing so, it draws on Delphi studies (n = 7) undertaken elsewhere and, where appropriate, these studies are used to illustrate potential options around key areas. Specific areas for consideration are: aims and objectives, sample selection, data collection, analysis and statistical interpretation, credibility, reliability and validity, and ethical issues.

Overview of Delphi technique

Delphi as a research methodology has been variously presented as a survey (Wang et al., 2003), procedure (Rogers and Lopez, 2002), method (Linstone and Turoff, 1975; Crisp et al., 1997) and technique (Broomfield and Humphries, 2001; Snyder-Halpern, 2002; Sharkey and Sharples, 2001). In this study, we refer to Delphi as a 'technique' because this appears to be the most commonly used terminology in the research literature. This technique is named after the ancient Greek oracle at Delphi who offered visions of the future to those who sought advice (Gupta and Clarke, 1996, p.185). There is general agreement that it was first used in technology forecasting studies initiated by the RAND (Research and Development) Corporation for the American military in 1944 (Gupta and Clarke, 1996). Since that time, it has become a popular way of engaging opinion from people with expertise, although, the technique itself and the purposes for which

it has been used have been extensively modified by researchers over the years (Gupta and Clarke, 1996; Crisp et al., 1997).

Most authors draw on all (Wang et al., 2003) or some (Gupta and Clarke, 1996; Robertson and MacKinnon, 2002) of the definition set out by Linstone and Turoff (1975) who define the Delphi technique as:

> *a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem.*

> (Linstone and Turoff, 1975, p.3)

In this study, on the development of a national set of child well-being indicators we are adopting the definition presented by Linstone and Turoff (1975).

Types of Delphi

A number of different types of 'Delphi' studies have been identified. Indeed, Gupta and Clarke (1996) note that 'practitioners are often willing, and sometimes even eager, to modify Delphi' (Gupta and Clarke, 1996, p.189). They conclude that while some modifications are useful, others are random and undermine both the quality and credibility of the technique. Hasson et al (2000) report that 'Modified Delphi', 'Policy Delphi' and 'Real-time Delphi' have all been used, although the following categorisation, described by van Zolingen and Klaassen (2003), has broad appeal:

The Classical Delphi: This type of study is characterised by five features. They are: anonymity, iteration, controlled feedback, statistical group response and stability in responses among those with expertise on a specific issue. Participants in this type of Delphi have expertise and give opinions to arrive at stability in responses on specific issues.

The Policy Delphi: The aim in this situation is not to reach stability in responses among those with expertise but to generate policy alternatives by using a structured public dialogue. Here the Delphi is an instrument for policy development and promoting participation by obtaining as many divergent opinions as possible. It is characterised by 'selective anonymity', iteration, controlled feedback, polarised group response and structured conflict. Selective anonymity may mean that participants answer questions individually but may also come together in a group meeting.

The Decision Delphi: This type of Delphi is used for decision making on social developments. Reality is created by a group of decision-makers rather than from the ad-hoc decision of only a small number of persons. Crucial to this type is that decision-makers involved in the problem participate in the Delphi. They are selected according to their position in the hierarchy of decision-makers and the aim is to structure thinking so that consensus can be achieved. The characteristic is 'quasi-anonymity' (where people with expertise are mentioned by name and known to everybody from the beginning but questionnaire responses are anonymous).

In this study we are using the Delphi technique to structure thinking around areas of child well-being so that consensus can be achieved in respect of a national set of well-being indicators. The proposed study will, therefore, adopt the approach of a 'Decision' Delphi.

Paradigmatic assumptions underpinning the Delphi technique

Some consideration is given here to the assumptions underpinning the Delphi technique. On first examination, it appears that the epistemological basis for the Delphi technique favours the positivist paradigm. Such a paradigm assumes the position of the researcher within the research to be that of an objective and uninvolved observer (Robson, 1993) and it could be argued that this is the case for the Delphi technique. The objectivist position in the Delphi technique is

supported through the utilisation of a quantitative approach to data collection and the application of single statistical measures to the identification of 'consensus'. The inclusion of 'experts' assumes an ontological position of single reality (on which 'experts' agree) and the reductionist approach to the identification of the phenomenon under study could also be understood as adhering to positivistic principles (Blackburn, 1999; Monti and Tingen, 1999). Others, however, present the Delphi technique as subjective and qualitative in nature (Fitzsimmons and Fitzsimmons, 2001) and we present the case for this understanding below.

The aim of employing a Delphi technique is to achieve consensus through a process of iteration. The process itself is concerned with opinions, ideas and words (Stewart, 2001) and it is suggested here that the purpose of the methodology (to achieve consensus through group interaction) is in keeping with an interpretative paradigm. Group interaction in research is generally underpinned by an assumption that an individual's attitudes and beliefs do not form in a vacuum and that people need to listen to others' attitudes and understandings so that they can focus on their own (Marshall and Rossman, 1995; Reed and Roskell, 1997). Within an interpretative paradigm, there can be many differing paradigms including, for example, post-positivism, critical theory, constructivism and participatory/co-operative paradigm (Lincoln and Guba, 2000). Within this, constructivism, and particularly social constructivism, appears to have most to offer in terms of understanding the epistemological basis for the Delphi technique. Lincoln and Guba (1985), writing about constructivism as a research paradigm, for example, note that:

> *Researchers in a variety of disciplines in the social sciences have been and are grappling with social constructivist approaches wherein the contribution of each individual in the context to the creation of a reality is recognised.*

> (Lincoln and Guba, 1985, p. 82)

A key advantage of the Delphi technique is the potential it holds to recognise and acknowledge the contribution of each participant and this is central to our study. While Schwandt (2000) writes that we are all constructivists if we believe that the mind is active in the construction of knowledge, social constructivists generally subscribe to an exogenic tradition of knowledge. Here the focus 'is on the arrangement of environmental inputs necessary to build up the internal representation' rather than on the person's intrinsic capacities for reason, logic or conceptual processing (Gergen, 1995, p.18). Within the Delphi technique, a process of individual feedback about group opinion, with opportunities for respondents to change their position, primarily on the basis of that feedback, provides a close fit with the use of environmental inputs to build up internal representations.

On the basis of the foregoing discussion, it is difficult to draw clear conclusions about paradigmatic assumptions underpinning all Delphi studies, since it is reasonably clear that certain parts of the technique are more coherent with a constructivist paradigm and others more coherent with that of positivism. In this study, we are suggesting that social constructivism has something to offer to our understanding of Delphi and we are, therefore, underpinning our study with features closely associated with that paradigm. Specifically, we are seeking to achieve individual reconstructions that coalesce around consensus through providing opportunities for knowledgeable participants to interact with each other in a structured way. We are also committed to an ontological assumption that there are multiple realities.

Purpose of Delphi technique

The main purpose of adopting a Delphi technique to decision-making is to provide a structured approach to collecting data in situations where the only available alternative may be an anecdotal or an entirely subjective approach (Broomfield and Humphries, 2001). The features of anonymity, iteration with controlled feedback, statistical group response and expert input can facilitate

consensus where there is contradictory or insufficient information to make effective decisions (Linstone and Turoff, 1975; Goodman, 1987; Hasson et al., 2000; Snyder-Halpern, 2002). Other group approaches to reaching consensus have been examined but have been found to be less appropriate to the development of a set of well-being indicators. These include, for example, nominal groups (Carney et al., 1996), brain-storming (Hasson et al., 2000), focus groups (Morgan, 1997) and analytic hierarchy process (AHP) technique (Lai et al., 2002) as well as the establishment of working groups. The main disadvantages with each of these techniques is their risk of taking account only of the perceptions of the most outspoken or opinionated members of that group or of only focussing on interesting or controversial elements (Fein et al., 1997).

Delphi technique as a methodology has been in use for almost sixty years and the types of situations where it can be useful have been well described. Although some methodological issues remain outstanding, it is noted that the Delphi technique has been found to be particularly useful in the following situations:

1) where a problem does not permit the application of precise analytical techniques but can benefit from subjective judgements on a collective basis;
2) where the relevant specialists are in different fields and occupations and not in direct communication;
3) where the number of specialists is too large to effectively interact in a face-to-face exchange and too little time and/or funds are available to organise group meetings; and
4) where ethical or social dilemmas dominate economic or technical ones (Linstone and Turoff, 1975; Gupta and Clarke, 1996).

In the proposed study on developing indicators of child well-being, understandings may be significantly influenced by the particular area of expertise, experience or occupational position of a participant. There is a significant benefit, therefore, in being able to harness subjective judgements of

respondents. Delphi technique has been found to be an appropriate mechanism for ensuring that emergent differences between and within key stakeholder groups (that can arise from differences in focus, situation and context) can be accounted for in a systematic way. Indeed, it is difficult to envisage a more appropriate way in which communication between the multiplicity of stakeholders concerned with children's well-being can be facilitated. The number of relevant people with expertise appears to be too large to interact in a meaningful way through face-to-face exchange, despite the very strong case for representation of as high a number of views as possible.

Advantages of employing a Delphi technique

An overarching aim of the Delphi technique is to achieve consensus and, while further discussion of its conceptual basis takes place later in this paper, we understood consensus here to mean 'a general agreement' (Thompson, 1995) 'in constructing a hierarchy and making judgements' (Lai et al., 2002, p. 135). It has been reported that the capacity of a group to reach consensus is influenced by:

- a tendency of low status group members to 'go along with' the opinion of group members with a higher status;
- a tendency of a group to exert pressure on its members to conform; and
- a tendency of a dominant group member to exert undue influence on the opinion of the group (Gupta and Clarke, 1996; Fein et al., 1997; van Zolingen and Klaassen, 2003, p. 318).

Arising from these, the most obvious advantage of guaranteed anonymity in responding to individual questions is that it is likely to encourage opinions that are free of influences from others and is therefore more likely to be 'true' (Goodman, 1987; Snyder-Halpern, 2002). It has been suggested that anonymity encourages experts to make statements on the basis of their personal knowledge and

experience, rather than a more 'cautious institutional position' (Gupta and Clarke, 1996, p. 186). By adopting an iterative approach to data collection through questionnaires and feedback however, the 'collective human intelligence capability' found in groups of people with expertise can be harnessed (Linstone and Turoff, 1975).

Other advantages relate to the use of questionnaires that have the capacity to capture a wide range of inter-related variables and multi-dimensional features (Gupta and Clarke, 1996, p. 186) and enable a geographically dispersed group of experts to provide their understandings (Rogers and Lopez, 2002). Respondents can complete the questionnaire at their leisure and this reduces time pressures and allows for more reflection and contemplation of response (Linstone and Turoff, 1975). This, in turn, may increase the number and quality of contributions and can decrease respondent burden by allowing participation at the participant's convenience. Snyder-Halpern (2002) summarises the 'primary' advantages of the Delphi technique as:

> *(i)ts adaptability to diverse data collection strategies, decreased peer pressure secondary to anonymity and the ease of condensing opinions of many and varied experts into a few precise statements.*

(Snyder-Halpern, 2002, p. 185)

Disadvantages of adopting a Delphi technique

Disadvantages of the Delphic technique have also been identified and authors have questioned the reliability, validity and credibility of this research methodology. Sackman (1975), for example, has noted that anonymity may lead to a lack of accountability because responses may not be traced back to the individual. In addition, it has been suggested that a consensus approach can lead to a diluted version of the best opinion and the result represents the 'lowest common denominator' (Powell, 2003, p. 378). It could be argued,

however, that all approaches (for example, working groups, nominal groups) to gaining consensus run this risk. Others have argued that this approach is time-consuming, labour intensive and, therefore, expensive (Fitzsimmons and Fitzsimmons, 2001) although there is not agreement about this (Powell, 2003). A number of methodological issues arising in respect of Delphi have the capacity to threaten the credibility of the study and these include issues around panel expertise, number of rounds, questionnaire development, analysis and achievement of consensus. These issues are considered in detail in the next section of this paper having particular regard to our study.

Summary

To summarise, the Delphi technique is a research approach to gaining consensus through the utilisation of questionnaires and the provision of feedback to participants who have expertise in key areas. While there are many potential types of Delphi techniques, three broad categories are generally in use and these are: Classical, Policy and Decision. In the proposed study we will most closely follow the approach of a 'decision Delphi' . The advantages and disadvantages of the Delphi technique are complementary. The adoption of an anonymous approach to data collection can facilitate positional openness and at the same time may lead to a risk of lack of accountability. The potential to harness a wide variety of views about different variables and across geographical areas has greater appeal than other alternatives, which may involve face-to-face interaction, but in doing so some nuances may be lost. On balance, however, we consider the advantages of this approach to outweigh the disadvantages for our particular situation. Some consideration has been given to the epistemological basis of the Delphi technique and while the literature around this is conflicting, we are suggesting here that many of its assumptions in terms of purpose and process are in keeping with that of social constructivism. Consequently, we understand the 'decision Delphi' technique to be primarily situated within an interpretative paradigm. Key issues relating to the

proposed study are now set out, using material from Delphi studies undertaken elsewhere on indicator development.

Proposed Study

This section sets out key issues relating to Delphi as a research technique and takes account of the aim and objectives, conceptual framework and key methodological issues emerging. In particular, it focusses on the identification of expertise, number of rounds, consensus, questionnaire development and data collection, analysis, and mechanisms for ensuring credibility in the proposed study.

Seven key studies are used throughout this section to illustrate options where contentious issues arise. These studies were identified in a systematic way using the principles of mini-review set out by Griffiths et al. (2002). The papers were identified through database searches of Medline (1966-2003), Embase (1980 - 2003), Cinahl (1982-2003), Eric (1988-2003), Social Work Abstracts (1977-2003) and Econlit (1988-2003) using two key words, 'Delphi' and 'indicator'. Because of the large number of abstracts generated (more than 3,000 in one database), it was necessary to place limitations on the material and these are identified below. Criteria for inclusion were that the study:

1. focussed only on indicator development;

2. reported on the research itself;

3. was published in English; and

4. was published after 1994.

A bibliographic review published in 1995 (Gupta and Clarke, 1996) identified that the three most popular areas for Delphi applications were health, education and business. Consequently, in identifying studies for illustrative purposes, we have taken random examples from health (Cambell et al., 2000; Wang et al., 2003; Schuster et al., 1997), education (Rogers and Lopez, 2002; van Zolingen and Klaassen, 2003) and business (Miller, 2001; Snyder-Halpern, 2002). This has facilitated maximum variability across different disciplines, an important factor given the diverse backgrounds of the likely participants in this study.

We accept that the criteria used for final inclusion of studies has limitations in respect of the material identified but we are, nevertheless, reasonably confident that the papers identified are sufficiently comprehensive to provide an overview of key areas for consideration. Table 2.1 sets out the main elements of each of the key studies.

Table 2.1: Overview of key studies

	Author	Main focus	Focus
1	Campbell et al. (2000)	'Prescribing' indicators for general practice	Health
2	Rogers and Lopez (2002)	Cross-Cultural school psychology competencies	Education
3	Schuster et al. (1997)	Quality of health care measures for children and adolescents	Health
4	van Zolingen and Klaassen (2003)	Key qualifications in senior secondary vocational education	Education
5	Snyder-Halpern (2002)	Indicators of organisational readiness for clinical information technology innovation	Organisational development
6	Millar (2001)	Indicators for sustainable tourism	Organisational focus
7	Wang et al. (2003)	Reproductive health indicators for China's rural areas	Health

Aim and Objectives

The National Children's Strategy (2000) identifies a set of actions emerging from Goal no 2: 'Children's lives will be better understood'. The key action to which this study relates is the development of a set of 'child well-being indicators' as a basis for the production of a bi-annual report to be known as the 'State of the Nation's Children'. The development of the set of well-being indicators is underpinned by the commitment within the strategy to the 'whole child perspective'. The aim of this study is, therefore:

To reach consensus about a national set of child well-being indicators that can be used as the basis for the bi-annual report 'the State of the Nation's Children.

The objectives of this study are, as follows:

- to gain consensus about indicators that take account of key aspects of the whole child perspective as set out in the National Children's Strategy;
- to gain consensus about indicators that will facilitate comparisons between the Irish and international context regarding child well-being;
- to gain consensus about indicators that meet key quality criteria.

Conceptual underpinning: the whole child perspective

The whole child perspective is used to provide a conceptual underpinning for this study. This perspective is informed by the work of Bronfenbrenner (1979) and Ward (1999) and is underpinned by the key principles of non-discrimination, best interests of the child, survival and development and respect for the voice of the child that emerge from the UN Convention on the Rights of the Child. The 'whole child perspective', first identified in the National Children's Strategy, takes account of the innate capacity of the child as well as the broader socio-ecological aspects of the environment. It sets out three broad domains and these are:

1. child's own capacity;
2. children's relationships; and
3. formal and informal supports.

Children's own capacity: The National Children's Strategy (2000) identifies nine dimensions of childhood and these are named:

- physical and mental well-being;
- emotional and behavioural well-being;

- intellectual capacity;

- spiritual and moral well-being;

- identity;

- self care;

- family relationships;

- social and peer relationships; and

- social presentation.

The National Children's Strategy (2000) notes that it is helpful to consider the outcomes children achieve at each stage of development as 'expressions', which develop over time and which eventually provide the capacity for coping with adulthood (National Children's Strategy, 2000, p. 25).

Any single dimension, for example, physical and mental well-being, presented above could form a legitimate focus for indicator development on children. Here, however, our intention is to focus on the Irish children in a holistic way and consequently, in addition to understanding individual dimensions we are also concerned with how different dimensions interact with each other and how they may influence and are, in turn, influenced by the other two substantive domains (children's relationships and formal and informal supports).

Children's Relationships: Children's relationships within the whole child perspective range from the family, which is acknowledged as the primary source of care, all the way up to the State, which acts as the ultimate guarantor of their rights. Within this, there may be some focus on 'within-family' relationships including, for example, the composition of families in which Irish children are born and in which they grow up.

Formal and informal supports: Essential services and supports, according to the whole child perspective, are provided through the primary social networks of family, extended family and community (informal supports), and through formal

support services provided by the voluntary sector, commercial sector, the State and its agencies.

The extent to which indicators about formal and informal supports will be identified as part of the indicator set on well-being is not pre-determined in this study, although members of the panel of expertise, will be asked to keep the three broad domains of the whole child perspective to the fore when making decisions about areas for indicator development. Work already undertaken in the Irish context highlights the difficulties in reaching a definition of child well-being being (Carroll 2002). For the purposes of guiding this study, however, we define child well-being as:

> 'Healthy and successful individual functioning (involving physiological, psychological and behavioural levels of organisation), positive social relationships (with family members, peers, adult caregivers, and community and societal institutions, for instance, school and faith and civic organisations), and a social ecology that provides safety (e.g., freedom from interpersonal violence, war and crime), human and civil rights, social justice and participation in civil society' (Andrew et al., 2002).

Figure 1. presents a graphic representation of the whole child perspective where the importance of the active developing child at the core of the model is noted.

Figure 1: Whole child perspective (National Children's Strategy 2000 p. 26)

Methodological Issues emerging

This section of the paper outlines the study method and focusses, in particular, on issues relating to sampling, data collection, analysis and statistical interpretation, credibility, reliability and validity, and ethical issues of the proposed study.

Sampling: Panel expertise

Delphi's claim to credibility lies in its ability to draw on expertise (Miller, 2001) and this is promoted by purposeful selection of 'experts' for inclusion to the panel, rather than relying on random sampling. The term 'expert' is contested (Hasson et al., 2000) and it has been suggested that this title is misleading (McKenna,

1994). Cognisant of this debate, we are proposing to use the terms 'panel of expertise' and 'participants' in this study rather than 'experts'.

There is some agreement that key features of participants involved in Delphi studies include both 'willingness' and 'ability' to make a valid contribution to the subject under examination (Goodman, 1987). These two factors need to be balanced with the potential for bias. Rowe et al., cited in van Zolingen and Klaassen (2003), suggest that researchers may create a study bias if they only:

a) select respondents who are easily available;
b) select respondents whose reputations are known to the researcher;
c) select respondents who meet a minimal number of criteria of familiarity with the field of the research problem; and
d) select respondents on the basis of self-ratings of their expertise.

An additional problem has been noted by van Zolingen and Klaassen (2003), who suggest that participants willing to take part in the Delphi method may be more favourable to the method. This, in turn, may mean they are more inclined to agree with other panel members than those having a less adaptable attitude. Consideration is now given to ways in which other research studies have determined eligibility for inclusion on panels of expertise and these are summarised in Table 2.2.

Table 2.2: Key issues relating to expert panels

	Author	N: Panel size	Expert eligibility
1	Campbell et al. (2000)	N = 305	All pharmaceutical and medical advisors in the UK
2	Rogers and Lopez (2002)	N = 65 including Practitioners, faculty and administrators	**Two of five criteria:** 1. Authorship 2. Conference presentation 3. Member of chair of committee 4. Employed in practice or supervision with five years' experience 5. Employed as faculty member with specific interest in area
3	Schuster et al. (1997)	N = 18 Two panels 9 members on each	**Paediatric panel: Nominations** from various academies - all physicians **Women's panel: nominations** from various academies - all physicians Representation from different clinical practice settings, community and academic medicine and geographic regions
4	van Zolingen and Klaassen (2003)	N = 53 4 stakeholder groups	Two years' **experience** in their function Extensive **knowledge** of the content of the functions, qualification, connections and carers For teachers: Had worked on curriculum development committee Senior teacher
5	Snyder-Halpern (2002)	N = 34	**Volunteers** identified through online membership directory of specific group Employed in US-based healthcare organisation Held position that reflected direct involvement with clinical information systems
6	Millar (2001)	N = 74	**'Informed academics and consultants'** 1. Had published in the area of sustainability in pervious two years in one of 4 major journals
7	Wang et al. (2003)	N = 123 Two panels Chinese (n = 63) International experts (60)	No information on Chinese panel Others identified through Ford Foundation reproductive health programme. **Background:** public health, social science, health care and women's studies. Had rural experience

The presentation in Table 2.2 illustrates the many diverse ways in which 'expert' can be defined, and mechanisms for identification of participants have ranged from 'volunteers' to 'nominations' to acknowledgement of 'experience and

knowledge'. Exact and explicit criteria are set for inclusion in the panel for some studies (Rogers and Lopez, 2002) while, for others, 'expertness' is assumed on the basis of membership of a particular group (Campbell et al., 2000) or organisation (Snyder-Halpern, 2002). In the case of the study undertaken by Schuster et al. (1997), nominations were sought from various academies of paediatrics and adolescent medicine (Panel 1) and family physicians and obstetricians (panel 2). Rogers and Lopez (2002) included practitioners, faculty and administrators (n = 65) in their study of developing indicators of cross-cultural school psychology competencies. They set out very explicit selection criteria for inclusion in their study as 'expert'. These criteria are identified in the example below, which referred to school psychologists who had expertise in the provision of psychological services to racially, culturally and/or linguistically diverse populations. The expertise of the panel members was defined in terms of professional accomplishments in multiple domains of professional functioning relevant to psychological service delivery with diverse clients, and two of the following five criteria had to be met.

- primary or secondary author of two or more school psychology publications concerning racially, ethnically, culturally and linguistically diverse clients;
- presented three or more presentations on relevant cross-cultural topics at national school psychology conferences;
- member or chair of an APA division 16 (school psychology) or NASP committee about delivering services to racially, ethnically, culturally and linguistically diverse clients;
- employed as a practising or supervising school psychologist with at least five years' experience working primarily with racially, ethnically, culturally and linguistically diverse clients; and
- employed as a school psychology faculty member at a school psychology program that emphasised multicultural or bilingual training.

In other studies, less formal criteria were identified and a snowball type approach was adopted. Wang et al. (2003), for example, report 'the majority of the Chinese panel experts were identified by the study team, while most of the international experts were chosen with the help of the Ford Foundation reproductive health program officers' (Wang et al., 2003, p. 218).

Number of participants and panels

There is no precise mechanism for identifying the number of individuals or the number of panels for inclusion in any individual study (Williams and Webb, 1994). It has been suggested that the size of the panel may vary according to the topics covered, the nature of different viewpoints included, and the time and money available (van Zolingen and Klaassen, 2003). One or more panels can be formed. Van Zolingen and Klaassen (2003), for example, included four different stakeholder groups in one panel, while Wang et al. (2003) included two panels differentiated by location (Chinese experts n = 63; international experts n = 60). Schuster et al. (1997) formed two panels (although each panel included only nine experts). In other studies, single panels were formed, although there was substantial variance in the numbers of experts included. Campbell et al. (2000), for example, included more than three hundred experts in their study, while Rogers and Lopez (2002) included sixty-five. There does not, therefore, appear to be an optimum number of panels or indeed, panel members.

In the context of the current study, the multi-dimensional nature of the whole child perspective, coupled with a desire to create a 'national' set of indicators, means that two options arise in respect of the number of expert panels. These are:

1. that a single heterogeneous panel of expertise be formed or,
2. that a number of separate panels be developed and that these be differentiated by, for example, background, focus or orientation.

Advantages and disadvantages of adopting one or more panels are set out in Table 2.

Table 2. Advantages and disadvantages of single and multiple panel

Advantages multiple panel	Disadvantages multiple panel
- The pool of expertise for any individual area would be larger - A more inclusive approach to stakeholders could be adopted and this could be advantageous in terms of 'buy-in'. - Stakeholders would only be involved in identifying indicators where they have specific areas of expertise	- May be conceptually inappropriate because the unifying feature of the 'whole child perspective' is the underlying commitment to understanding children in a holistic way - The breadth of focus of any individual panel may be difficult to determine - The complexity of the study would increase exponentially - Anonymity may be difficult to maintain - The balance between selecting experts who would be relatively impartial and yet have information that reflects current knowledge may be difficult to strike - It is possible that, having initially adopted a reductionist approach, the complexity and number of 'experts' involved in different panels may lead to several practical problems in the subsequent integration of the material
Advantages single panel	Disadvantages single panel
- It would be conceptually more coherent - The identification of each indicator would be situated within the context of the overall 'whole child perspective' and this may have a synergetic quality. - The study would be less complex than that of multiple panels - The focus of the development would be on the 'whole child perspective' and consequently it may be easier to strike the balance between impartiality and expertise	- It may not be possible to include the same number in the panel, so the subsequent level of buy-in may be lower - Panellists may respond to areas where they do not have expertise and this may create difficulties in interpretation - It may not be possible to reach consensus because the diversity of the panel experts may be too great

Panel of expertise for this study

Of critical importance to the development of a national set of child well-being indicators is that children be understood in a holistic way. We are, therefore, committed to reaching consensus across many different areas of children's lives and, consequently, it seems logical to have a single panel of expertise. Such an approach will protect against fragmentation and lack of coherence within the indicator set. It also, however, raises some issues around the extent to which specific indicators within the indicator set should be determined by all panel members. In order to accommodate this, we intend to use the Delphi technique to identify the broad areas for inclusion in the indicator set, rather than the individual indicators themselves.

Selection of panel experts for this study

The selection of panel experts is fundamental to the credibility of the study and two main possibilities for the identification of participants, one employing systematic sampling and the other snowball sampling, are set out below. Five broad constituency groups (categories) who hold expertise in the area of child well-being have been identified and these are:

1. children;
2. statutory policy-makers and service providers;
3. non-government organisations with a focus on child well-being;
4 researchers and others with experience in specific areas not covered by category 2 or 3; and
5 parents.

Children: In keeping with Goal no. 1 of the National Children's Strategy (2000), *'children will have a voice in matters which affect them and their views will be given due weight in accordance with their age and maturity'* (National Children's Strategy, 2000, p. 30) we have considered ways in which children can be included in this process. We have concluded, following discussion, that the most appropriate way in which this can be done is by commissioning a study on

'children's understandings of well-being'. This study will adopt a group approach to:

- identify, collate and analyse themes emerging from group discussions with children about their understandings of well-being;
- identify ways, from children's perspective, in which these themes can best be represented; and
- examine the extent to which these themes are commensurate with the whole child perspective as set out in the National Children's Strategy.

This study is currently underway and is being directed by Dr Saoirse Nic Gabhainn, National University of Ireland, Galway. We believe it will be possible to link the findings from that study with the Delphi study using the process of triangulation. Triangulation can be supported here because theoretically, as we have set out above, we understand this study to be underpinned by social constructivism, where the influence of social interaction in building up an internal representation is important (Gergen, 1995). This is also reflected in the theoretical underpinning of group approaches to data collection.

Systematic sampling

The potential for systematic sampling with the remaining stakeholder groups has been given consideration. We have examined ways in which, for example, statutory policy-makers and service providers can be identified, particularly those in the main departments working with children's issues. We have found, however, that although there are six main departments working with children's issues, there are many sub-divisions within each. Within the Department of Health and Children, for example, there are more than fifty separate organisations (including the National Children's Office) as well as the main department. Although a full list of each of these was compiled, with a view to creating a sampling frame, it was clear that such an approach was unwieldy, would lead to confusion, and was unlikely to result in a panel of expertise.

Problems also arose in respect of the research community because of the absence of a central database in the Republic of Ireland for academics or researchers with an interest in children's issues. Nevertheless, we did make efforts to identify those who had published in the area of children's research through, for example, individual institutional databases that identify publications from researchers / lecturers working in their university and other on-line portals, for example, 'Expertise Ireland'. Again, we concluded that the likelihood of achieving a comprehensive sampling frame would be beyond the scope of this study. Within the non-governmental organisations there is also considerable expertise in respect of children's issues. The Children's Rights Alliance (CRA), for example, brings together more than seventy organisations which work with, or on behalf of, children. In light of the difficulties highlighted above, however, and in view of some concerns about the ability to capture the expertise of those working on a day-to-day basis with children, it was decided that this type of systematic approach to sampling was not an option.

Snowball sampling

As a workable alternative to the complexities outlined above, other options were examined, and in identifying a panel of expertise we have adopted a snowball sampling approach. We were greatly facilitated in this by the Children's Research Advisory And Development Group, set up by the National Children's Office in 2003. This group comprises eighteen people and members represent various Government departments (including Education and Science, Social, Community and Family Affairs, Justice, Equality and Law Reform and Health and Children), research communities and service providers from a multiplicity of different areas including sociology, education, social policy, health and law.

Each member of the advisory group will be invited to take part in the study themselves and will also be asked to nominate two other people. In addition, the Children's Rights Alliance will also be invited to take part in the study and to nominate others with expertise. Using this approach, it is expected that the overall composition of the group will reflect the diversity of children's lives. We

expect that the final panel will comprise approximately eighty people, with expertise in different areas and working in different contexts.

Parents: Some consideration has been given to the inclusion of parents in the panel of expertise. There is no structure, however, for identifying a sampling frame of parents. Various organisations, such as the 'National Parents Council' have been contacted and nominations to the study sought. Other mechanisms for identifying parents are also being examined.

Data collection

Notoriously low response rates for questionnaire surveys can be minimised by ensuring that respondents are fully informed about the study and that reminders are issued (Robson, 1993; Cohen and Manion, 1994). Some consideration has been given here to whether panel participants could be brought together prior to distributing the first questionnaire. This would have some benefits in terms of providing information for participants about the importance of continuing to engage with the study, and would also facilitate the presentation of key issues relating to the study. This, however, would not be in keeping with the Dephi technique and for that reason we do not intend to bring people together at this point in the study. We will, however, ensure adherence throughout the course of the study to good practices in relation to maintaining response rates.

Other questions arise in respect of how the questionnaires should be distributed. The ease, convenience and comprehensiveness of e-mail may provide a possibility for on-going involvement of respondents. Advantages and disadvantages as well as key issues emerging in respect of this, identified by Snyder-Halpern et al. (2003) are set out in Table 2.4.

Table 2.4 Advantages and disadvantages for e-mail use

Advantages	Disadvantages
Recruitment was cheaper and quicker	Disadvantages mainly technological and include:
Cost savings were substantial	Unexpected changes of e-mail application
Legibility of responses	Unreliability of panellist e-mail capabilities, resulting in some
Ease of data entry, resulting in decreased data entry time and errors	panellists being unable to participate in some rounds
Decrease in response turnaround time	Some respondents were unable to retrieve e-mail attachments in their original format
Ability to track transmission status (for example, for incorrect addresses)	Some respondents had incompatibility with Excel or MS word applications.
	Problems with mime encryption

(Snyder-Halpern, Thompson C. and Schaffer J. (2003) 'Comparison of mailed vs. internet applications of the Delphi Technique' in *Clinical Informatics Research*

The authors, based on their experiences, have made a number of recommendations to ensure good response rates and these are: 1) all files should be saved down to an earlier version of Microsoft Word and Excel prior to mailing; 2) the complete round questionnaire should be sent with each reminder mailing to avoid time delays due to requests for an additional questionnaire copy; and 3) it may be necessary to establish a separate e-mail account for managing study data.

In addition to e-mail contact, we intend to make the questionnaire available for completion on-line and by post. This is being facilitated by the use of Key-Point © software and is supported by the computer department at the Department of Health and Children.

Round 1

In the classical Delphi technique the first round adopts an inductive approach, where participants are invited to generate ideas and are given complete freedom in their responses (Hasson et al., 2000). This allows panel members to identify an infinite range of possibilities, and although this has many advantages in terms of comprehensiveness, it can be problematic where the respondents are heterogeneous in knowledge, expertise and experience. In addition, this

inductive-type approach has been criticised on the basis of its inability to produce the level of information that a thorough literature review would produce (Millar, 2001). In this study, we have considered the appropriateness of an open-ended approach to the first round but believe that, on balance, it is more efficient to adopt a semi-structured approach for the following reasons: a substantial amount of work has already taken place in respect of indicator areas and it is important that this work would inform the development of a national set. Two sequential processes in respect of indicator sets have taken place and these are:

First, a systematic search for indicator sets commonly used elsewhere was undertaken, and more than 1,300 indicators were identified (Brooks, 2003). Although some of these indicators were clearly of less relevance to the Irish situation (for example, an indicator of 'percentage of children carrying guns to school'), and some were almost identical to each other, it was decided not to exclude any indicators at that point.

Second, an expert in the area of data sources for child statistics (Fitzgerald, 2003) was contracted by the National Children's Office to examine the extent to which data were available for each of the indicators identified by Brooks (2003). These indicators were then categorised according to whether data sources were, or were not, currently available, and this material is now with the National Children's Office. In doing so, we have a better understanding of the sources which may be available to us in respect of many different areas. It is important to note here that we are not using the current availability of data sources as an exclusion or inclusion criterion at this point.

In addition to the work already undertaken, a key objective of this study is 'to identify indicators that will facilitate comparisons between the Irish and international context regarding child well-being'. It is likely, therefore, that the final indicator set will contain some indicators that will allow for such

comparability. Consequently, it is more efficient to provide information to participants about indicator areas in use elsewhere.

On the basis of the work already completed, we are proposing that the first round questionnaire should set out broad areas in an 'event list' and that participants be asked to rate the most important areas to children's well-being for inclusion in the indicator set. This will then be used to guide the development of the second and third questionnaires as appropriate. This type of approach has been used by others (for example, Schuster et al., 1997; Millar, 2001; Rogers and Lopez, 2002; van Zolingen and Klaassen, 2003) and it has been reported that an 'event list', similar to that being proposed, is more preferable than a blank piece of paper because it provided participants with a context in which to provide their responses (Snyder-Halpern, 2002).

Questionnaire development

In developing questionnaires for each round, we will adhere to good practices and will be particularly mindful of issues raised about the length of the questionnaire, the importance of clearly formulated unambiguous questions and the implications of having a heterogeneous group of respondents (van Zolingen and Klaassen, 2003). The initial questionnaire will be divided into two parts. Part one will ask respondents to identify their demographic and 'expertise' characteristics, while Part two will concern itself with the broad areas for indicator development.

Issues to which we have given some consideration include:

- number of indicators to be included: in general, indicator sets have between 25 and 35 indicators and it is our intention to approximate to this number;

- question content: this has been informed by the extensive review of indicators already undertaken and will take account of the whole child perspective and the definition set out by the Ecology Working Group (2002) on child well-being;

- question wording: the questionnaire will take the form of an event list and this will get over difficulties relating to question wording;

- form of response to the question: some consideration has been given to whether a 'rating' or 'ranking' approach should be taken. We have opted for the use of a rating scale (from 1- 10, where 1 is least important and 10 is most important) because we believe there will be too many broad areas initially to rank (approximately 60). In addition, the adoption of a rating approach allows for the use of measures of dispersion (e.g. standard deviation) which will be central to the identification of stability between rounds;

- place of the question in the sequence: we have taken the view that responses will be alphabetically sequenced.

Cohen and Manion (1994) caution against the use of leading questions, 'highbrow' questions (even with sophisticated respondents), complex questions, irritating questions, and questions that use negatives and this will be a consideration in this study. 'Subletting questions' by using alphabetical symbols (for example, Q1a) or numerical symbols (for example, Q2(1)), repeating instructions, putting ticks in boxes, and coloured paper as well as special attention to typography (including inter-line spacing, headings, font size) will all be considered so that the questionnaire will be as attractive and easy to respond to as possible (Orna and Stevens, 1995). Issues arising in respect of construct, content and face validity of the questionnaire will be considered and this will be done by pre-testing and piloting.

Pilot study

The situation around pilot testing prior to the main study is unclear. Powell (2003), for example, argues that 'pilot testing is optional although it may be useful to

identify ambiguities and improve the feasibility of administration' (Powell, 2003, p. 378). In view of the importance of this study to future policy-making and developments around children's lives, we have undertaken substantial pre-testing of the first round questionnaire. We have also carried out a two-round pilot study with ten participants, and these participants will not be involved in the final study. The results highlighted a number of issues of importance and these have been addressed in the first round questionnaire for the main study. The pilot study also facilitated a link with the study on 'children's understandings of well-being' and both research teams met on completion of their respective pilot studies.

Potential questionnaire format for the identification of specific indicators

Some authors have reported extensively on their questionnaire (Millar, 2001) while others have only indicated broad areas. Miller (2001), for example, presents a clear outline of key questions asked in their study of indicators for sustainable tourism and this is set out in Figure 2.2.

Figure 2.2: Questionnaire format for indicators (Millar 2001)

Name of indicator: _____

Type of indicator: _____

Page number (in explanatory booklet): _

		Yes	No	See notes
1	Is the indicator applicable?			
2	Is the indicator a complete indicator?			
3	Is the indicator applicable to all types of tourism?			
4	Is the data for the indicator easily obtained?			
5	Is the calculation required for the indicator simple?			
6	Is the indicator understandable?			
7	Is the data objective, quantifiable and reliable?			
8	Does the indicator point towards sustainable development?			
9	Can the indicator be measured on an ongoing basis?			

Others report using a smaller number of questions for each indicator. Campbell et al. (2000), for example, in a study of quality indicators, report three questions for each indicator. Respondents in that study were asked 'to rate each indicator against two continuous 1-9 integer scales', 'Is this indicator a useful measure of cost minimization' and 'Is this indicator a useful measure of quality'. Respondents were also asked whether they currently used each indicator. The questionnaire also invited comments from respondents.

Figure 2.3: Example of question format (Campbell et al. 2000)

Example indicator format in round 2 questionnaire showing the three different types of feedback	Cost minimisation		Quality	
Antibiotic generic prescribing rate (%)	49 26 32 22 11 5 4 1 2 9 8 7 6 5 4 3 2 1	7	19 12 26 27 15 15 13 10 16 9 8 7 6 5 4 3 2 1	6

Summary of comments:
Some comments advocated a high %—for example, "high generic prescribers are sometimes those using the older established antibiotics more than the new high cost ones." Other comments focused on why a high % may not be appropriate—for example, "expensive 2nd and 3rd line drugs are often prescribed generically but inappropriately."

Snyder-Halpern (2002) reported asking three questions for each sub-dimension. The first employed a four-point likert-type scale (where 1 = not important and 4 = critically important), the second asked whether the dimension should be retained, modified or deleted and the final question asked about how, if applicable, the dimension should be modified. This is similar to that presented by Schuster et al. (1997), although in that study a nine- rather than four-point likert-type scale was used.

The heterogeneous nature of the expert group in this study may mean that respondents will not have expertise in each indicator area. One potential way of taking this into account is to use the information provided in the demographic / expertise section to include or exclude respondents. This subjective approach may, however, lead to problems of interpretation. The most common approach appears to be either to ask participants to identify their level of expertise for each question or, alternatively, to respond only to questions where they have

expertise. Munier and Rondé (2001, p. 1543), for example, asked respondents to declare their knowledge on a given subject based on four clearly defined levels (very good, good, limited and null) using very precise definitions set out below:

- *'Very good'* knowledge implied that the 'expert currently devotes himself to research on this precise topic or closely related topic';
- *'Good knowledge'* suggests that the 'expert devoted himself in the past to this research and continues to follow very closely the work of his colleagues';
- *'Limited knowledge'* means the 'expert is satisfied with reading articles, newspapers and reviews or has contacts with specialists'; and
- *'Null knowledge'* suggests that the 'expert is not informed in the field'.

An analysis of responses by all fields (78,486) showed that only 3% reported having 'very good' knowledge, 7% having 'good' knowledge and almost two-thirds (63%) had 'null' knowledge. This, in turn, may have serious implications for the number of respondents for each indicator.

In this study, we are of the view that all participants on the panel of expertise have some knowledge of the broad areas for inclusion, by virtue of their involvement in children's lives. More specific detail will be required when identifying the actual indicator areas, and at that point only those with specific knowledge in the area will be involved. This will accommodate the need for different areas of expertise when identifying the specific detail of the individual indicator.

Consensus

Consensus has been identified as one of the most contentious components of the Delphi method, and debates have centred on the position of consensus in the overall study. The aim of the Delphi technique is to achieve consensus but this is not a straightforward concept and is generally poorly explained in studies (Williams and Webb, 1994). The *Longman Dictionary of Contemporary English*

(Thompson, 1987) defines consensus as 'a general agreement; the opinion of most people in a group' (*Longman Dictionary of Contemporary English*, 1987, p. 216). Although some authors have presented qualitative judgements of consensus (e.g. Millar, 2001), in general an empirical approach is taken. Consensus is usually determined through statistically measuring the variance in responses across rounds. Less variance is understood to mean greater consensus (Rowe and Wright, 1999) although this has itself been the subject of some controversy. Bardecki (1984), cited in Rowe and Wright (1999), reported that respondents with more extreme views were more likely to drop out of the study than participants with more moderate views. The conclusion drawn was that the decrease in variance can be a consequence of attrition rather than consensus.

Munier and Rondé (2001), among others, suggest that the possibility that participants may simply alter their estimates in order to conform to the group (conformance), without actually changing their opinions (consensus), must be considered. Their own work in testing the influence of expert knowledge on consensus suggests, however, that consensus is the more likely explanation for decreased variance. Their conclusion that 'it can be theoretically demonstrated that the median response of the entire group should move towards the true value' supports a move towards consensus rather than conformance (Munier and Rondé, 2001, p. 1548). Some authors have undertaken 'post group consensus', which concerns the extent to which individuals on completion of the Delphi technique, agreed with the final group aggregate, their own final round estimates or the estimates of other participants (Rowe and Wright, 1999).

Studies focussing on the number of rounds needed in a Delphi survey to achieve consensus suggest that most changes occur in the transition from the first to the second round (van Zolingen and Klaassen, 2003). The number of rounds in the modified technique may be decreased to as few as two, if panellists have been provided with an event list, and if early group consensus is achieved (Snyder-Halpern, 2002). Other authors have focussed on participant burden as a problem and suggest that when the number of rounds exceeds four, the response rates

can be very low. Table 2.5 sets out definitions of consensus used in studies elsewhere.

Iteration is a key feature of the Delphi technique and feedback on questionnaire analysis is provided to each respondent at each round. Feedback has been defined as:

'The means by which information is passed between panellists so that individual judgement may be improved and debiasing may occur.

(Rowe and Wright, 1999, p. 370)

Levels of feedback vary and may be provided in a number of different ways. The purpose of feedback is to allow each expert to revise his or her own judgement in light of the judgement of others (Munier and Rondé, 2001). Consequently, it has been suggested that in addition to asking experts to provide a statistical summary, two additional pieces of information should be provided. These are the average self-rated expertise of all the experts and also the reasons why particular scores are provided. This would allow respondents to place their own responses in the context of others' level of expertise as well as their rationale. Others have suggested that feedback be provided in the form of a median or inter-quartile range, and that experts who continue to give extreme views are asked to provide a rationale (van Zolingen and Klaassen, 2003). Crisp et al. (1997) notes that one of the most common forms of feedback is measures of central tendency (mean, median), which may or may not be accompanied by a measure of dispersion (standard deviation).

The timing of feedback is also an issue and it has been suggested that the quality of the Delphi study increases as the time between filling in a questionnaire and the next one being mailed becomes shorter (Waldron, cited in van Zolingen and Klaassen, 2003).

We intend to provide individualised feedback according to indicator area so that members of the panel of expertise will be able to situate their responses within the broader context.

Analysis

Analysis that takes place in a Delphi study has two purposes. First, analysis should provide feedback between rounds for respondents and, second, it should be able to identify when consensus has been reached. There does not, however, appear to be agreement about the best method of mathematical aggregation (Murphy et al., 1998). In Rowe and Wright's (1999) systematic review of literature on Delphi technique, a number of different descriptive statistics were used. These included median, mode, percentages for each event, ranks, upper and lower quartile ranges, regression weights or induced (if-then) rules, statistical average of points for each factor. Qualitative material was also examined and in a number of studies 'reasons' were analysed and given to respondents as feedback.

Table 2.5 presents two variables (main statistics used and definition of consensus) reported on in a small number of studies used for illustrative purposes in this paper. In general, statistics presented included:

- percentages;
- standard deviations;
- means;
- medians; and
- ranges.

While cautioning that the above statistics assume an interval scale, Greatorex and Dexter (2000) concluded that the mean, a measure of central tendency, can be understood to represent group opinion. The standard deviation (a measure of spread), they suggest, can then be understood as a representation

of the amount of disagreement within the panel. If the standard deviation is low, then the panel is in agreement and the converse is also true. If the standard deviation is high, the panel is in disagreement. In a systematic examination of consensus development methods and their use in clinical guideline development, however, Murphy et al. (1998) argue that the median and the inter-quartile range are more robust than the mean and standard deviation.

Statistics used in the seven key studies are summarised in Table 2.5, along with the definitions of consensus.

Table 2.5 Analysis and definition of consensus in key studies

	Author	Main focus	Statistics	Consensus
1	Campbell et al. (2000)	'Prescribing' indicators for general practice	Percentages for each indicator	Disagreement defined as 30% or more scores in both the bottom (1-3) and top (6-9) tertile
2	Rogers and Lopez (2002)	Cross-cultural school psychology competencies	Item means and standard deviations Median Range	'The percentage of panellist ratings that fell within the established range of consensus +/- 1.64SD on round 2' (p. 127)
3	Schuster et al. (1997)	Quality of health care measures for children and adolescents	Number of votes in each category (1-9) for each indicator	Consensus not required because each panellist has a vote for each indicator (p1086). Disagreement defined as at least 3 votes in the 1-3 range and at least 3 votes in the 7-9 range
4	van Zolingen and Klaassen (2003)	Key qualifications in senior secondary vocational education	Information file - no additional information	Qualitative understanding: 'opinions did not differ in many cases' (p. 337)
5	Snyder-Halpern (2002)	Indicators of organisational readiness for clinical information technology innovation	Seven member coding team in thematic analysis of responses	Inter-relater agreement = 43% or greater i.e. 3/7 or greater
6	Millar (2001)	Indicators for sustainable tourism	Percentage agree / disagree Mean Standard deviation	Qualitative: using words like 'general agreement', 'disharmony', 'Spread of opinion', 'disagreement', 'Divisive'
7	Wang et al. (2003)	Reproductive health indicators for China's rural areas		

In this study, we will use the median as the main statistical measure and this will be particularly useful in facilitating a reduction in the number of areas for inclusion in the indicator set. The cut-off level will be determined by relative rating and it is likely, for example, that the top twenty-five areas according to the median will be included in the indicator set and those below will be excluded. This approach will be used in conjunction with the standard deviation which will provide a measure of dispersion.

Credibility of the study

Since Sackman's (1975) initial critique of the Delphi technique and Linstone and Turoff's (1975) claim that this method is more 'art than science', a substantial literature has developed about the credibility of Delphi as a research method. Criticisms of the Delphi technique eloquently summed up by Gupta and Clarke (1996) include:

> *Conceptual and methodological inadequacies, potential for sloppy execution, crudely designed questionnaires, poor choice of experts, unreliable result analysis, limited value of feedback and consensus, and instability of responses among consecutive Delphi rounds.*

> (Gupta and Clarke, 1996, p. 187)

In essence, there is potential for compromising credibility at all stages of the study. Some authors have argued that the Delphi technique is an interpretative research approach and should be judged accordingly (Keeney et al., 2001). This, however, is problematic because, although there is some agreement that the terminology and criteria used to judge qualitative enquiry should differ from that of quantitative studies, there is little agreement around the precise criteria that should be used (Emden and Sandelowski, 1999; Lincoln and Guba, 2000;

Cutcliffe and McKenna, 2002). According to Powell (2003), a decision trail should be set in a way that provides sufficient evidence to defend the appropriateness of the method to address the problem selected, choice of expert panel, data collection procedures, identification of justifiable consensus levels and means of dissemination and implementation

Some of these areas have also been identified by others, operating from a more positivist paradigm (e.g. van Zolingen and Klaassen, 2003). They have suggested that compromises to the validity and reliability of the study arise from the value-led nature of feedback and consensus and the instability of responses. These areas are, in turn, influenced by the number of experts, their average expertise and the average inter-correlation of their judgements. These authors argue for setting specific guidelines around each of these areas, so that the reliability of the study (or whether a replication of the study would give the same results with a different panel) can be judged.

Others have suggested that the credibility of the study rests on whether it is effective in aiding decision-making. The extent to which this is the case has been subject to some examination. One systematic review of empirical studies (n = 25) comparing Delphi study with standard interacting groups concluded, with some caution, that Delphi groups outperform groups in decision-making and forecasting (Rowe and Wright, 1999).

Throughout this paper, we have identified areas of potential compromise in the carrying out of this study and have presented a literature on areas around which methodological decisions need to be taken. We view this as the conception of a decision trail and suggest that responses and decisions taken on the basis of questions raised can provide a mechanism for making judgements about the credibility of the study.

In addition, we have engaged with a small group of advisors, each with a different area of expertise pertinent to the study. These advisors are:

Dr Ed. Carroll: Author, National Youth Council of Ireland

Ms Deirdre Cullen: Central Statistics Office

Ms Eithne Fitzgerald: National Disability Authority

Dr Maeve Henchion: Teagasc

This group has provided guidance and support for us while, at the same time, providing a potential monitoring mechanism to ensure transparency in the conduct of the study. We believe this further enhances the credibility of this study.

Ethical issues

Ethical issues 'saturate all stages of the research process' and start with the researcher's choice of topic and method (Punch, 1998, p. 281). We believe it is ethical to adopt a Delphi approach to identify a national set of child well-being indicators. First, such an approach will facilitate the engagement of more expertise than any other group method. In addition, this type of study facilitates 'fair' representation of the views of each participant because each participant has an equal opportunity to have their views taken into account. Alternative mechanisms for reaching consensus do not provide as transparent a decision trail for each indicator, and the capacity of the Delphi technique to achieve this, means that the rationale for inclusion and exclusion of indicators can be clearly stated. This is likely to lead to greater acceptance of the findings than other methods.

The potential for harm in this study is relatively low, because participants will be mature adults and, as each will be chosen on the basis of their expertise, they are not considered vulnerable. Nevertheless, other ethical issues revolving around consent, privacy and confidentiality of data will also be considered. It is our intention that informed consent will be achieved at each stage of the study.

Participants will be informed about the purpose of the study, the procedures to be followed, the anticipated time commitment, and contact details for the principal investigators if they wish to ask any questions about the study. They are, of course, free to withdraw from the study at any time.

In the research context, the right to privacy can be violated during the course of an investigation or after the study has been completed. In this case, every effort will be made to protect the privacy of the participants. Two ways of protecting privacy are through confidentiality and anonymity. The essence of anonymity is that information provided by participants should in no way reveal their identity (Cohen et al., 2000) and such anonymity is a central feature of the Delphi technique. Individual names or positions will not be directly linked to individual responses in the questionnaire feedback.

There is an expectation by participants in almost all studies that confidentiality will be protected and this is also the case in this study. Assurances of confidentiality will be given to all participants and, at the onset, a code number will be allocated to each participant. Completed questionnaires will be identifiable only by code number and the key for the code will be held in a locked filing cabinet. This will be accessible only to members of the advisory group. Questionnaires and other data collected will be held in a secure location for a period of ten years after the study but will then be destroyed. Requirements under data protection legislation will be complied with.

In summary, this study will actively subscribe to principles of mutual respect, non-coercion and non-manipulation, the support of democratic values, and the belief that every research act implies moral and ethical decisions (Denzin and Lincoln, 1994). These principles will be used to guide each part of the study and at all stages, issues relating to consent, privacy and confidentiality will be key features.

Conclusion

To conclude, this paper has set out key issues for the proposed methodology to develop a national set of child well-being indicators in Ireland. Seven studies on indicator development using the Delphi technique have been used to illustrate key issues around the proposed study. These issues include those related to sampling, panel expertise, questionnaire development for the initial round and additional rounds, data collection and analysis. Key aspects of consensus, validity, reliability and credibility as well as ethical considerations have been made explicit, and ways in which each will be dealt with identified. Figure 2.4 summarises key elements of the proposed study in terms of panel composition and expertise, number of rounds and outcomes and provides a graphic overview of how elements of the study may proceed. Appendix 1 presents a time frame for the proposed study.

<u>Figure 2.4 Key elements of proposed study</u>
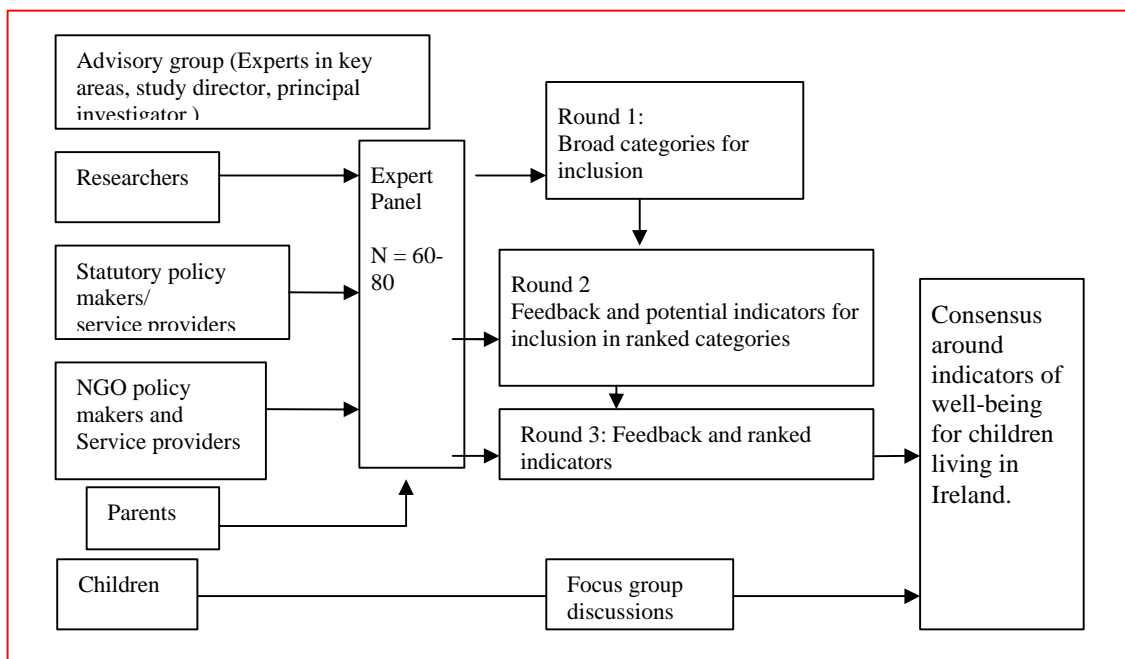
References:

Blackburn, S. (1999) Think. Oxford: Oxford University Press.

Bronfenbrenner, U. (1979) The ecology of human development. Cambridge, Ma: Harvard University Press.

Brooks, A.M. (2003) 'Compilation of national and international child well-being indicators'. Dublin, National Children's Office.

Broomfield, D. and Humphries, G.M. (2001) 'Using the Delphi technique to identify the cancer education requirements of general practitioners', *Medical Education* 35 928-937.

Campbell, S.M., Cantrill, J.A. and Roberts, D. (2000) 'Prescribing indicators for UK general practice: Delphi consultation study', *BMJ* 321 425-428.

Carney, O., McIntosh, J. and Worth, A. (1996) 'The use of the nominal group technique in research with community nurses', *Journal of Advanced Nursing* 23, 5, 1024-1029.

Carroll, E. (2002) The Well-being of Children: Four Papers Exploring Conceptual, Ethical and Measurement Issues Dublin: Irish Youth Foundation.

Cohen, L. and Manion, L. (1994) Research Methods in Education. 4th Edition. London: Routledge Falmer.

Cohen, L., Manion. L. and Morrison, K. (2000) Research Methods in Education. 5th Edition. London: Routledge Falmer.

Crisp, J., Pelletier, D., Duffield, C., Adams, A. and Nagy, S. (1997) 'The Delphi method?', *Nursing Research* 46(2) 116-118.

Cutcliffe, J.R. and McKenna, H.P., (2002) 'When do we know that we know? Considering the truth of research findings and the craft of qualitative research', *International Journal of Nursing Studies* 39 611-618.

Denzin, N.K., and Lincoln, Y.S. (1994) 'Introduction: Entering the field of qualitative research', in: Denzin, N.K., and Lincoln, Y.S. (eds.)   Handbook of Qualitative Research. London: Sage: 1-17.

Ecology working group (2002) Ecology of Child Well-Being: Advancing the Science and the Science-Practice Link Georgia: Centre for Child Well-Being.

Emden, C., and Sandelowski, M. (1999) 'The good, the bad and the relative, part two: Goodness and the criterion problem in qualitative research', *International Journal of Nursing Practice* 5 2-7.

Fein, J.A., Lavelle, J.M., Ginsburgt, K.R. and Giardino, A.P. (1997) 'A methodology to maximize resident input in improving a pediatric rotation' *Arch. Pediatr. Adolesc. Med.* 151 840-844.

Fitzgerald, E. (2003) 'A feasibility study on sources of child well-being indicators', Dublin, National Children's Office.

Fitzsimmons, J.A. and Fitzsimmons, M.J. (eds.) (2001) Service management: Operations, Strategy and Information Technology (4th Edition) Boston, McGraw-Hill .

Gergen, K.J. (1995) 'Social construction and the educational process' in: Steffe, L. and Gale, J. (eds.) Constructivism in Education. Hove, UK: Lawrence Erlbaum Associates: 17-40.

Goodman, C.M. (1987) 'The Delphi technique: a critique', *Journal of Advanced Nursing* 12 729-734.

Greatorex, J. and Dexter, T. (2000) 'An accessible analytical approach for investigating what happens between the rounds of a Delphi study', *Journal of Advanced Nursing* 32(4) 1016-1024.

Griffiths, P., 2002. 'Introducing the mini review', *British Journal of Community Nursing* 7(1), 38.

Guba, E.G. and Lincoln, Y.S. (1994) 'Competing paradigms in qualitative research' in: Denzin, N.K. and Lincoln, Y.S. (eds.) <u>Handbook of Qualitative Research</u>. London: Sage: 105-117.

Gupta, U.G. and Clarke, R.E. (1996) 'Theory and Applications of the Delphi Technique: A bibliography (1975-1994)', *Technological Forecasting and Social Change* 53, 185-211.

Hasson, F., Keeney, S. and McKenna, H. (2000) 'Research guidelines for the Delphi survey technique', *Journal of Advanced Nursing* 32(4) 1008-1015.

Keeney, S., Hasson, F. and McKenna, H.P. (2001) 'A critical review of the Delphi technique as a research methodology for nursing', *International Journal of Nursing Studies* 38 195-200.

Lai, V.S., Wong, B.K. and Cehung, W. (2002) 'Group decision making in a multiple criteria environment: A case using the AHP in software selection', *European Journal of Operational Research* 137 134-144.

Lincoln, Y.S. and Guba, E.G. (1985) <u>Naturalistic inquiry</u>. London: Sage.

Lincoln, Y.S. and Guba, E.G. (2000) 'Paradigmatic controversies, contradictions, and emerging confluences', in: Denzin, N.K. and Lincoln, Y.S. (eds.) Handbook of qualitative research. London: Sage: 163-188.

Linstone, H.A. and Turoff, M. (eds.) (1975) The Delphi Method Techniques and Applications. Massachusetts, Reading: Addison-Wesley.

McKenna, H. (1994) 'The Delphi technique: a worthwhile research approach for nursing?', *Journal of Advanced Nursing* 19 1221-1225.

Marshal, I.C. and Rossman, G.B. (1995) Designing qualitative research. 2nd Edition. London: Sage.

Millar, G. (2001) 'The development of indicators for sustainable tourism: results of a Delphi survey of tourism researchers', *Tourism Management* 22 351-362.

Morgan, D.L. (1997) Focus Groups as Qualitative Research. 2nd Edition. London: Sage.

Monti, E.J. and Tingen, M.S. (1999) 'Multiple paradigms of nursing science', *Advances in Nursing Science* 21(4) 64-80.

Munier, F. and Rondé P. (2001) 'The role of knowledge codification in the emergence of consensus under uncertainty: empirical analysis and policy implications', *Research Policy* 20 1537-1551 .

Murphy, M.K., Black, N.A., Lamping, D.L., McKee, C.M., Sanderson, C.D.B., Askham, J. and Marteau, T. (1998) *Consensus development methods, and their use in clinical guideline development: a review* United Kingdom, Health Technology   Assessment   2(3)

http://www.hta.nhsweb.nhs.uk/pdfexecs/summ203.pdf accessed September 13th 2003

National Children's Strategy (2000) *The National Children's Strategy – our children- their lives* Dublin, Government publications.

Orna, E. and Stevens, G. (1995) <u>Managing information for research</u>. Buckingham: Open University Press.

Punch, K. (1998) <u>Introduction to social research</u>. London: Sage.

Powell, C. (2003) 'The Delphi technique: myths and realities', *Journal of Advanced Nursing* 41(4) 376-382.

Reed, J. and Roskell, V. (1997) 'Focus groups: issues of analysis and interpretation', *Journal of Advanced Nursing* 26 765-771.

Robertson, H.A. and MacKinnon, N.J. (2002) 'Development of a list of consensus-approved clinical indicators of preventable drug-related morbidity in older adults', *Clinical Therapeutics* 24(10) 1595-1613.

Robson, C. (1993) <u>Real World Research</u>. Oxford: Blackwell.

Rogers, M.R. and Lopez, E.C. (2002) 'Identifying critical cross-cultural school psychology competencies', *Journal of school psychology* 40(2) 115-141.

Rowe, G. and Wright, G. (1999) 'The Delphi technique as a forecasting tool: issues and analysis', *International Journal of Forecasting* 15 353-375.

Sackman, H. (1975) <u>Delphi critique</u> Lexington Books, Massachusetts, Lexington.

Schuster, M.A., Asch, S.M., McGlynn, E.A., Kerr, E., Hardy, A.M. and Gifford, D.S. (1997) 'Development of a Quality of Care measurement system for children and adolescents: Methodological considerations and comparisons with a system for adult women', *Arch. Pediatr. Adolesc. Med.* 151 1085-1092.

Schwandt, T.A. (2000) 'Three epistemological stances for qualitative inquiry' in: Denzin, N.K. and Lincoln, Y.S. (eds.) Handbook of Qualitative Research. 2nd Edition. London: Sage: 189-213.

Sharkey, S.B. and Sharples, A.Y. (2001) 'An approach to consensus building using the Delphi technique: developing a learning resource in mental health', *Nurse Education Today* 21 398-408.

Snyder-Halpern, R. (2002) 'Indicators of organizational readiness for clinical information technology/systems innovation: a Delphi study', *International Journal of Medical Informatics* 63 179-204.

Snyder-Halpern, R., Thompson, C. and Schaffer, J. (2003) 'Comparison of mailed vs. internet applications of the Delphi Technique in Clinical Informatics Research' http://www.amia.org/pubs/symposia/D200120.PDF Accessed 18th July 2003.

Stewart, J. (2001) 'Is the Delphi technique a qualitative method', *Medical Education* 35 922-923.

Thompson, D. (1995) The concise Oxford dictionary of current English. Oxford: Clarendon Press.

van Zolingen, S.J. and Klaassen, C.A. (2003) 'Selection processes in a Delphi study about key qualifications in Senior Secondary Vocational Education', *Technological Forecasting and Social Change* 70 317-340.

Wang, C.C., Wang, Y., Zhang, K., Fang, J., Liu, W., Luo, S., Tang, S., Wang. S. and Lu, V.C. (2003) 'Reproductive health indicators for China's rural areas', *Social Science and Medicine* 57(2): 217-225.

Williams, P.L. and Webb, C. (1994) 'The Delphi technique: a methodological discussion', *Journal of Advanced Nursing* 19 180-186.