

The Delphi method

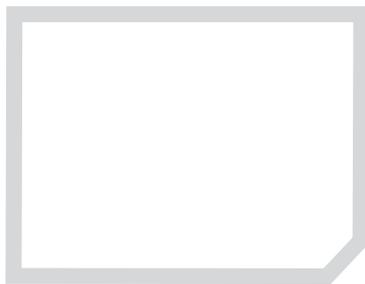
Susanne Iqbal and Laura Pison-Young with a step-by-step guide

The Delphi survey method is popular in many disciplines. Originally developed in the US as a means of forecasting future scenarios, this method has been used to determine the range of opinions on particular matters, to test questions of policy or clinical relevance, and to explore (or achieve) consensus on disputed topics. Although there is considerable variation in how the method is applied, the Delphi method has its own distinct characteristics:

- I It uses a group of participants (known as 'panellists') specially selected for their particular expertise on a topic.
- I It is often conducted across a series of two or more sequential questionnaires known as 'rounds'. It employs an initial 'idea generation' stage, in which panellists

are asked to identify the range of salient issues.

- I It collates ideas from Round 1 to construct the survey instrument distributed in subsequent rounds.
- I It has an evaluation phase (third or further rounds) where panellists are provided with the panel's responses and asked to re-evaluate their original responses.
- I It is interested in the formation or exploration of consensus, often defined as the number of panellists agreeing with each other on questionnaire items.



Measuring diversity or steering towards consensus?

The Delphi method is particularly useful in areas of limited research, since survey instruments and ideas are generated from a knowledgeable participant pool (Hasson et al., 2000), and it is suited to explore areas where

controversy, debate or a lack of clarity exist.

However, the many applications and descriptions available in the literature can be confusing, and the Delphi method, with a few exceptions, remains relatively unexploited in psychological research (e.g. Graham & Milne, 2003; Haggard & Haste, 1986; Haste et al., 2001; Jeffery et al., 2000; Petry et al., 2007). To make the Delphi method more accessible for psychologists unfamiliar with this method, this article provides a practical step-by-step guide based on our experiences of conducting Delphi studies in clinical psychology. It is not an exhaustive account, and further guidance is available elsewhere (e.g. Hackett et al., 2006, Graham & Milne, 2003, Keeney et al., 2006, Schneider & Dutton, 2002).

What is the aim of your study?

The first step is to determine whether the study aims to measure the diversity of opinions on a topic or to steer a group towards consensus. This is an important distinction in terms of the execution of the Delphi. In general, if your study aims to generate consensus, three or more rounds are preferable. Ideally, the same panel should be retained throughout and high response rates are particularly important in order to determine the impact of group feedback on panellists.

On the other hand, if the Delphi process is a means of measuring opinions, fewer rounds are generally acceptable. Having a complete dataset is less vital, and the panel can be expanded across rounds by inviting more panellists in Round 2.

Strongly disagree	1	2	3	4	5	6	7	8	9	10	Strongly agree	Can't say		Your choice was
Overall percentages	0	0	0	4	0	0	4	13	29	50		0	100% N=24	10

Figure 1: Fictitious example of an individualised feedback in Round 3 statement: Smoking should be banned in hospitals.

references

Beretta, R. (1996). A critical review of the Delphi technique. *Nurse Researcher*, 3(4), 79-89.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77-101.

De Meyrick, J. (2003). The Delphi method and health research. *Health Education*, 103(1), 7-16

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

Graham, L. & Milne, D. (2003). Developing basic training programmes: A case study illustration using the Delphi method in clinical psychology. *Clinical Psychology and Psychotherapy*, 10, 55-63.

Graneheim, U. & Lundman, B. (2004). Qualitative content analysis in nursing research. *Nurse Education Today*, 24, 105-112.

Hackett, S., Masson, H. & Philipps, S. (2006). Exploring consensus in practice with youth are sexually abusive: Findings from a Delphi study of practitioner views in the United Kingdom and the Republic of Ireland. *Child Maltreatment*, 11(2), 146-156.

Haggard, M. & Haste, H. (1986). One generation after 1984: Psychology in the year 2010. *Bulletin of the British Psychological Society*, 39, 321-324.

Hardy, J.D., O'Brien, A.P., Gaskin, C.J. et al. (2004). Practical application of the Delphi technique in a bicultural mental health nursing study in New Zealand. *Journal of Advanced Nursing*, 46(1), 95-109.

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

Decide the structure

The next step is to decide the number of rounds, to draw up a timeframe and to construct study materials (e.g. letters to participants, consent forms, complete ethics procedures).

A two-round Delphi (e.g. Petry et al., 2007) is most suitable when there is a clear literature base from which to establish the survey instrument and if the main aim is to take the temperature of opinion on a topic. Although quantitative questionnaires have been used in the first round, a qualitative first round is optimal, because the primary function of the Delphi method is to explore an area of future thinking that goes beyond the currently known or believed. Also, the reliability and validity of the study may be improved if an initial group of experts produces the items.

When exploring consensus, rounds may continue until consensus is reached. However, this approach can quickly compromise panellists' response rates and enthusiasm. Three rounds, which would typically take four months, often suffice (Stone Fish & Busby, 2005).

Selecting panellists

Panellists form the lynchpin of the Delphi, and clear inclusion criteria should be applied and outlined as a means of evaluating the results and establishing the study's potential relevance to other settings and populations. The number of panellists depends very much on the topic area as well as the time and resources at the researchers' disposal. Although Delphi surveys have been conducted with as few as seven and as many as 1000 panellists, Turoff (2002) recommends panels between 10 and 50. These numbers seem more appropriate, given the amount of data and subsequent analyses each panellist generates.

Researchers must also decide how to conceptualise and define 'expertise'. The method may be undermined if panellists are recruited who lack specialist

The Delphi in use

Iqbal et al. (in press) used the Delphi to explore and approach consensus in a study exploring sexually inappropriate behaviours in children under the age of 10. Review of the pertinent literature had revealed that children's sexual behaviours were judged differently by different professionals. The Delphi seemed appropriate to explore this sensitive topic as it is an ideal tool to expose all the different positions, including arguments for and against these positions, to generate consensus and to communicate this.

The Delphi is 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 & Turoff, 2002, p.3), and based on the idea that it is possible and valuable to reach a consensus (Stone Fish & Busby, 2005). By feeding back percentages of all views to each participant and inviting them to reflect on their responses in the light of these scores, consensus was achieved that children who display sexually inappropriate behaviours should not be called 'sex offenders'. No consensus was achieved with regard to many other sexual behaviours, particularly those considered 'normal'.

In line with the Delphi methodology, results were disseminated via participants themselves, reports, and journal publications. 'Divergence' (when no consensus was achieved) was also fed back to highlight how little agreement existed amongst very experienced professionals with regard to what counts as normal sexual behaviours. This should improve practice and allow for further research to be carried out.

knowledge, qualifications and proven track records in the field (Keeney et al., 2001), although of course expertise comes in many guises and may include those who are 'experts by experience' (Hardy et al., 2004). In general, a varied panel is considered best in producing a credible questionnaire, and individuals who might provide a minority or differing perspective should be actively recruited to the panel (Linstone & Turoff, 2002).

With regard to the recruitment process itself, panellists are often recruited via letter or e-mail. Recruitment can be broadened through 'snowballing' (asking panellists to pass on invitations to other relevant individuals).

The questionnaires

The more open-ended the Round 1 questionnaire (Q1) the better, ideally involving a series of open-ended questions inviting panellists to brainstorm. A quantitative 'tick-box' style format may also be used, but since the

Delphi method sets out to generate new ideas, a quantitative Q1 seems to defy this purpose.

The Q1 is usually created following a detailed literature review, consultation with relevant individuals and consideration of the aims of your Delphi study. Generally speaking, asking panellists to spend 30 minutes completing the questionnaire is considered reasonable, and pilot testing is essential to determine timeframes as well as readability and relevance of the questions.

Online surveys can be an efficient alternative to posting questionnaires and often appeal to panellists. Web services (e.g. surveymonkey.com) can be a simple way of constructing online questionnaires. Once the questionnaire has been distributed, following up non-responders is recommended as high response rates can improve the credibility of a study (Beretta, 1996). Ideally, a 70 per cent response rate should be maintained (Sumsion, 1998). We found that regular contact, flexibility around

Haste, H., Hogan, A. & Zachariou, Y. (2001). Back (again) to the future. *The Psychologist*, 14(1), 30-33.

Iqbal, S., Hackett, S. & Callanan, M. (in press). 'Normal' and 'inappropriate' childhood sexual behaviours: Findings from a Delphi study of professionals in the UK. *Journal of Sexual Aggression*.

Jeffery, D., Ley, A., Bennun, I. & McLaren, S. (2000). Delphi survey of opinion on

interventions, service principles and service organisation for severe mental illness and substance misuse problems. *Journal of Mental Health*, 9, 371-384.

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

Keeney, S., Hasson, F. & McKenna, H.

(2006). Consulting the oracle: Ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing*, 53(2), 205-212.

Likert, R. (1932). A technique for measurement of attitudes. *Archives of Psychology*, 140, 44-53.

Linstone, H.A. & Turoff, M. (Eds.) (2002). *The Delphi method: Techniques and applications*. Retrieved 20 May 2009

from <http://is.njit.edu/pubs/delphibook>.

Petry, K., Maes, B. & Vlaskamp, C. (2007). Operationalizing quality of life for people with profound multiple disabilities: A Delphi study. *Journal of Intellectual Disability Research*, 51(1), 334-349.

Sackman, H. (1975). *Delphi critique*. Boston: Lexington Books.

Schneider, J. & Dutton, J. (2002).

deadlines and individual ‘thank you’ messages increased response rates.

The Round 2 questionnaire is constructed from the data gathered from the Q1. Commonly, a quantitative, ‘tick-box’ style survey using Likert (1932) type agreement scales or ranking scales are used. The construction of the Round 2 questionnaire (Q2) is often time-consuming. The use of methodological tools such as qualitative content analysis (e.g. Graneheim & Lundman, 2004) or thematic analysis (e.g. Braun & Clarke, 2006) is necessary to make the study methodologically more robust. Furthermore, careful attention to principles of questionnaire design is vital, and extended piloting may be necessary to iron out ambiguous, repetitive or inaccurate items.

On return of the Q2, descriptive data analyses of the panel’s responses can begin so that the Round 3 questionnaire (Q3) can be constructed. The purpose of the Q3 is to invite panellists to consider their scores in the light of the group response and decide whether they want to change any of their responses. We suggest feeding back percentages and providing individual round scores for every item (see Figure 1). This provides a visual means for the panellists of assessing the diversity of responses. It also allows them to check that researchers have recorded correct responses.

Analyses and dissemination

Upon receipt of the completed Q3, you need to check whether any changes have been made, in which case the data need to be re-analysed. Percentages, medians, interquartile ranges, means and standard deviations are commonly calculated.

Results can be presented in various ways. This includes reporting only those items that have reached a pre-agreed level of consensus (e.g. Petry et al., 2007), listing all items in order of consensus magnitude (Hardy et al., 2004), or also reporting those areas in which there is debate amongst the panel.

Table 1: Advantages and disadvantages of the Delphi survey method

Advantages	Disadvantages
Very flexible methodology that can accommodate many variations and applications.	Method suffers from a lack of guidance and agreed standards regarding interpretation and analyses of results, universally agreed definitions of consensus, as well as criteria for how panellists should be selected (Sackman, 1975).
Draws together existing knowledge and pinpoints areas of agreement/disagreement.	Less efficient as a means of generating or testing new knowledge and theories.
Enables a group communication that otherwise might have been impossible due to geography, time or other constraints (Stone Fish & Osborn, 1992).	Generalisations are limited: another panel may reach different conclusions, and it cannot be concluded that the only or correct issues have been identified.
Economical in terms of financial outlay and participant time. Potentially rewarding research process for participants with multiple inbuilt opportunities for feedback.	High levels of commitment required from panellists; drop-out levels often high.
Makes the potentially confounding interpersonal processes often occurring in ‘live’ groups less likely (e.g. conformity to the dominant view).	May lack some of the richness and depth found in ‘live’ groups.
Anonymity between panellists can encourage creativity, honesty and balanced consideration of ideas (De Meyrick, 2003).	Anonymity may produce less ‘ownership’ of ideas. Delphi process assumes panellists are willing or able to elucidate issues individually and respond honestly.

Finally, disseminate your findings (write a consensus report, article, present findings to services, etc.) amongst concerned parties, including your participants.

Strengths and weaknesses

Like any other survey method, the Delphi method has strengths and weaknesses. These are summarised in Table 1, and further critiques can be found in

Goodman (1987), and Sackman (1975). In our experience, the benefits outweigh its drawbacks; and this method seems particularly relevant for psychology. Traditionally, there has been a divide between quantitative and qualitative methods. The Delphi method can straddle this divide. By virtue of its procedural structure (to incorporate both qualitative and quantitative methods), it provides the opportunity to achieve a more complete picture of the phenomenon under study.

Attitudes towards disabled staff and the effect of the national minimum wage. *Disability and Society*, 3, 283–306.
 Stone Fish, L. & Busby, D. (2005). The Delphi method. In D. Sprenkle & F. Piercy (Eds.) *Research methods in family therapy* (2nd edn, pp.238–253). New York: Guilford Press.
 Stone Fish, L. & Osborn, J. (1992). Therapists’ views of family life: A

Delphi study. *Family Relations*, 41, 409–416.
 Sumsion, T. (1998). The Delphi technique. *British Journal of Occupational Therapy*, 61(4), 153–156.
 Turoff, M. (2002). The Policy Delphi. In H. Linstone & M. Turoff (Eds.) *The Delphi method* (pp.80–96). Retrieved 22 May 2009 from <http://is.njit.edu/pubs/delphibook>.

I Susanne Iqbal is a Chartered Clinical Psychologist at George MacKenzie House, Fulbourn Hospital, Cambridge
 Susanne.Iqbal@cpft.nhs.uk

I Laura Pipon-Young is a Chartered Clinical Psychologist at the Women’s Service, Secure & Forensic Services, Hellingly, East Sussex

