How expert are the experts? An exploration of the concept of ‘expert’ within Delphi panel techniques

The use of the term ‘expert’ occurs widely in healthcare research, in the context of national guidelines and consensus methods for the development of clinical protocols. Within consensus methods of research, especially Delphi panel techniques, the use of ‘experts’ is fundamental to reliability. Yet literature fails to debate the practicalities of defining ‘experts’ for use within Delphi panel research. This paper, by John Baker and colleagues, draws on methodological literature and discusses the concepts and elements of ‘experts’. It concludes with recommendations for researchers to ensure rigor in selecting experts for future Delphi research.

Introduction
The term ‘expert opinion’ is widely used within the NHS. Expert opinion is commonly sought in the development of clinical protocols; for example, National Institute for Health and Clinical Excellence (NICE) guidelines and in the provision of evidence within inquiries into adverse incidents. The use of experts is a defining feature of consensus methods of research. This paper draws on methodological review papers and recently published Delphi panel...
research papers to explore the concept of the 'expert' within Delphi panel techniques. It discusses the inherent qualities required in defining an expert and how different types of Delphi panels and sampling techniques can influence the definition and therefore the choice of experts. The paper concludes by summarising the main elements required to improve the rigor and validity of the use of experts within Delphi research.

The development of the Delphi method has been attributed to Dalkey and Helmer (1963) of the RAND Corporation (Jeffery et al 2000, Keeney et al 2001, Mead and Moseley 2001, McBride et al 2003, Reid 1988), although the conceptual roots can be traced further back in time. It is reported that the concept originated in the legend of the Greek Delphi oracle, a Homeric poem to Apollo. The oracle utilised a number of informants to deliver the 'truth', enhanced as a result of data from many sources (Kennedy 2004). During the 1950s, the United States Air Force sponsored the 'Project Delphi', which was established to predict the outcome of a hypothetical Russian nuclear attack on the munitions output of the US (Dalkey and Helmer 1963). Dalkey and Helmer (1963) devised this methodology based on the notion that it would allow participants (n = 7) to make considered independent opinions leading to reliable conclusions. Following this, the technique became widely utilised within future forecasting. As a methodology it now has a 50-year history, emerging from north American usage in commerce and government to recent and widespread history in healthcare settings and social research, and is increasingly used by nurses (Beech 2001, Keeney et al 2001). A pivotal component of this type of research is the identification of a 'panel of experts'.

Since inception the reliance on experts within consensus research has been controversial. In the original studies there is no account of how or why experts were chosen, or the specific standards for selection of panelists (Dalkey and Helmer 1963). One of the panelists was so knowledgeable they also provided 'expert' advice to the researchers on the methodology. A major criticism has been a failure to account for the choice or definitions of experts. Indeed, Sackman's (1975) critique of the Delphi methodology listed the unconvincing definition of expert as one of ten major flaws in the Delphi method. In addition the quality of panellist has reduced over time. By 1975
panellists’ level of expertise was already reduced to the level of informed individuals (Linstone and Turoff 1975).

There continues to be a paucity of literature regarding the concept of experts (Mullen 2003, Walker and Selfe 1996). Crisp et al (1999) criticised researchers for the cursory attention they have paid to the concept of experts, suggesting that the concept has not been properly defined in the literature as a result. This has been further complicated as the classical Delphi has been adapted to include many hybrids such as ‘modified’, ‘realtime’ and ‘policy’ Delphis (Kenney et al 2001).

Why use experts?
The rationale for using Delphi techniques is clear: they form an established method for determining consensus on best policy (Beech 2001). In addition Graham et al (2003) state that the technique’s feasibility makes it ideal in areas where consensus is lacking, for treatment protocols and for other ‘best practices’ where agreement is desirable. Mead and Moseley (2001) suggest that it is particularly useful in areas of limited previous work, policy making or to develop practice guidelines, and Hardy et al (2004) state that it is ‘particularly useful when there is little knowledge or uncertainty surrounding the area being investigated’.

Experts provide an accessible source of information that can be quickly harnessed to gain opinion. They can often provide knowledge when more traditional research has not been undertaken. This, arguably, ensures high content, face-to-face and concurrent validity (Beech 2001, Sharkey and Sharples 2001).

Defining an ‘expert’
The dictionary definition of an expert is ‘a person who is very knowledgeable about or skilful in a particular area’ (Soanes and Stevenson 2003). Despite significant criticism in the literature over the last ten years about Delphi as a methodology, there remains little consensus as to who is an expert. Such a lack of clarity has resulted in wide variations in definitions (Keeney et al 2001, McKenna 1994, Reid 1998, Williams and Webb 1994). Parenté and Anderson-Parenté (1987) concluded that there were no guidelines for defin-
issues in research

ing an expert or evidence that using experts increased the accuracy of a Delphi study.

However, despite this there have been some attempts to define experts within the construct of Delphi studies. Mead and Moseley (2001) state that experts can be defined in a number of ways, such as their position in a hierarchy, public acknowledgement or as recommended by other participants in a study. Crisp et al (1999) suggest that the use of the word ‘expert’ may be inappropriate and suggest the term ‘informed advocates’ be used instead. This, they argue, is because few panels truly consist of experts. However, a critical review of the Delphi method by Keeney et al (2001) cites a range of definitions of ‘expert’ including ‘informed individual’, ‘specialist in the field’ or ‘someone who has knowledge about a specific subject’.

It would appear that there is limited consensus as to what an expert is. It may not be about who they are but what attributes they possess. An expert should be a representative of their professional group, with either sufficient expertise not to be disputed or the power required to instigate the findings (Fink et al 1984). However, too narrow a definition of expert reduces the potential sample size available (Duncan et al 2004). Key themes have emerged from these numerous definitions of expert, including knowledge and experience, and ability to influence policy (Cantrill et al 1996, Keeney et al 2001, Kennedy 2004). Further expansion of these key characteristics is vital in making an informed choice about expertise.

Knowledge
One characteristic through which experts are often defined is knowledge. This can manifest in many different attributes, such as a professional qualification or registration. The possession of a qualification means that an individual has achieved a certain predefined knowledge and experience base. Many authors cite a professional qualification in their definition of ‘expertise’ (Hardy et al 2004, Williams and Webb 1994). There are clear advantages in defining a level of knowledge, which should enable the researcher to have some consistency of knowledge within the panel. However, Crisp et al (1999) propose that registered qualifications are not consistent with expertise, drawing on the example of developing research priorities. They state that while a registered
nurse would be unable to define research priorities they would be able to identify areas where they have practical difficulties.

Indeed individuals can be in possession of knowledge without clinical experience. Keeney et al (2001) describe further criticism of the use of 'experts' who are defined by knowledge alone, suggesting that knowledge does not equal expertise. Knowledge can be demonstrated in ways other than a professional qualification: for example, possessing a higher degree in a specific area may increase the credibility of an expert. It would seem that an honours degree, with increasing and widening participation, should no longer be used as a defining point.

The authoring of materials such as books or peer-reviewed articles may demonstrate knowledge within an area and this has been utilised as a criterion for selecting an expert. Duncan et al (2004) selected individuals who had 'published treatment manuals or used them in published research' – this was the main criterion for their expertise. Graham et al (2003) selected participants on the basis of a minimum number of 'quality' papers published within the last three years with 'at least one paper in a peer-reviewed medical journal'. Other researchers have identified individuals that they feel are knowledgeable within the area (Jeffery et al 2000, McBride et al 2003, Mead and Moseley 2001, Philips 2000).

Care is needed, however, in ensuring that experts who are known personally to the researcher are not invited (Murphy et al 1998). This can cause difficulties when experts come from a small group of individuals who know each other. More controversial is to ask potential panellists to rate their own expertise (Mullen 2003). The effect on the results of those who rated themselves as high or expert compared to those who self-rated as non-expert needs to be established. Finally, Duffield (1989) discusses a case-weighting of expertness. This was, however, dismissed as an unworkable idea but it could prove invaluable for researchers to request potential experts' CVs in order to judge their expertise.

But a major criticism of the recruitment of experts is their potential for bias: participants with specific and cutting-edge knowledge in an area may have a vested interest in preventing research taking place or in manipulating the results (Keeney et al 2001). The expression of conflict of interest should be a
requirement of any potential panellists. Another danger in utilising experts in the development of clinical protocols is that those participating may possess knowledge but be very distant from clinical practice and therefore unable to articulate to practice their theory and/or knowledge (Sharkey and Sharples 2001). Caution is, however, required as experts should not be judged on their representativeness but on their quality (Powell 2003).

**Experience**

In many papers one inclusion criterion often cited to justify expertise is the establishment of a predetermined level of experience. This is often linked with a professional qualification (knowledge) and it is often specified that an individual should have worked within an area for a certain length of time (Hardy *et al* 2004, Jeffery *et al* 2000). Again, caution is clearly required: it is tenuous to suggest that a certain number of years’ experience means that an individual can be considered an expert. It may also be impossible to predict whether the individual will possess the necessary attitude, knowledge or skills if years of experience is the sole criterion upon which they are judged. No research has been identified that explores any evaluation of the nature of an individual’s experience and their resultant level of expertise. It is proposed that clinical practice may enable a healthcare professional to make valuable observations based on this experience. Evidence is currently unavailable as to whether working within a good organisation would provide better observations than working in a bad one.

However, experience as a criterion is important when applied to those other than ‘professional experts’. Delphi research often concentrates on professional (qualification and experience) expertise but clearly this does not fit with national policy on involving users of services. Indeed, having a professional qualification often precludes service user involvement, despite services users’ knowledge and experiences. The inclusion of patients or service users within an expert panel provides valuable insights and is undoubtedly important (Sumison 1998, Cantrill *et al* 1996, Fink *et al* 1984). Within the literature several authors have stated that service users have been included as experts if they have so many years of experience (Hardy *et al* 2004) or experience of, for example, an operation (Mead and Moseley 2001).
The inclusion of service users based on a predetermined number of years' experience is arguably difficult to justify. And it remains debatable – and controversial – whether service users can add additional validity to an expert panel dealing with technical information or expert opinion based on knowledge prerequisites. Care is required, however, as there is potential for service users' views to become marginalised very quickly as they may not share the same language of the professionals. Given that no two people share the same experience, if research is looking for commonality then this is harder to justify within a panel, although wide-ranging experience and viewpoints are essential in maximising the findings of a Delphi panel. Alternative methods of ensuring service users' viewpoints are encompassed include ratifying the findings, triangulating the results or utilising a different research modality. Arguably a minimum requirement would be to conduct further verification studies with service users to establish 'credibility' (Walker et al 2000).

Policy influence
A number of papers have cited positions such as nurse consultant or chief executive as part of a definition of expertise, or positions within key organisations including NICE, the Department of Health or pressure groups. Graham et al (2003) included 'opinion makers within national organisations' as a criterion for their study. Service users and carers are a vital component of any Delphi project aiming to target policy (Mullen 2003).

Homogenous or heterogeneous?
A major discussion within the literature is the debate about homogenous or heterogeneous samples. These appear to have a major influence on the resultant definitions of expert that researchers have utilised, as the two types of sample require very different sample sizes. In using a homogenous sample, a narrow definition of expert can be applied. Unfortunately this will reduce the potential sample size available (Duncan et al 2004). However, it could ensure that 'true' experts will be identified. The other extreme results in large, all-encompassing heterogeneous samples. The definition of 'experts' therefore influences the sample size necessary to ensure validity of the result.
Recent researchers have all suggested the need for heterogeneity of samples that include those from diverse settings (Hardy et al. 2004, Mead and Moseley 2001, Mullen 2003, Powell 2003). The belief that this approach is somehow better for the validity of the findings is widely cited (Mead and Moseley, 2001). Heterogeneous groups appear to be selected because if they agree then the findings must be worthwhile (Mead and Moseley 2001). However, the dangers of adopting this approach are rarely discussed. Agreement may be on the more trivial or non-relevant points because this is the only consensus the panel can reach. Additionally, the larger the samples the further from the original Delphi concept researchers stray. The original study was conducted with seven experts, and consensus suggests that the most reliable samples for Delphi studies should be small – fewer than 20 participants (Jeffery et al. 2000, Mullen 2003, Philips 2000). This is very difficult to achieve with a large heterogeneous sample.

**Panels may not need experts**

Sackman's (1975) major critique of Delphi panels maintains that expert and non-expert panels make little difference to outcomes, especially in relation to forecasting or evaluating social phenomena. It may be proposed that outcomes would be similar regardless of panel make-up. Two pieces of research have been identified that evaluate this claim. Walker (1994) made a direct comparison of two panels. One panel consisted of physiotherapist researchers and the other newly of qualified physiotherapists. Similar findings were reported between the two groups and the researcher concluded that the level of expertness required was uncertain. Secondly Duffield (1993) explored the responses of two comparable expert panels; 93 per cent were accepted or rejected by both panels. This, it is proposed, was indicative of the reliability of an expert panel regardless of participants.

**Follow-up of non-respondents**

As with other survey methodologies, if those that do not participate are different from the sample there is potential bias in the findings. Limited research has attempted to establish whether some experts are more likely to participate, whether there is a reason for this and what the effect is on the results. McKee et al (1991) questioned the representativeness of members
within expert panels. A sample of 503 doctors was selected: 246 (48.9 per cent) replied, and 166 (33 per cent) said they would participate. Those not willing to take part were asked why. The researchers found no significant difference between those who willing and unwilling to take part, in relation to time since qualification, specialty, sex, higher degree, or whether the doctor was a UK graduate. The only significant difference was that consultants with an appointment in a teaching hospital were less likely to take part. The authors suggested that the differences could be due to mailing factors. They concluded that expert panels were very similar in characteristics to their colleagues but were unable to identify further research in this area.

**Future research**
In order to increase the robustness of future research, defining the notion of expert is of vital importance. Work is also required to examine and refine selection criteria.

It could be proposed that there is a need for a consensus exercise to determine a hierarchy of expertness similar to present hierarchies of research. Perhaps expert panels could receive star ratings based on clear and consistent criteria. Hierarchy of language could accompany this so that not all panels are termed ‘expert’. How an expert is defined not only influences the make-up of a panel but also affects the sample size needed to make the research reliable. To the authors’ knowledge no research has compared a panel of professionals to one of service users. If panels are to be heterogeneous and include a diverse range of participants such undertakings are important. Likewise, it appears rare for researchers to compare directly or include members of different disciplines within the same panel.

**Conclusions**
This paper discusses the notion of expert within Delphi panel research. It is clear that experts are multi-faceted and there will continue to be difficulties in defining and justifying their selection. As Sumsion (1998) states: ‘Consideration of these options reveals that there is no ready answer and it becomes the responsibility of each researcher to choose the most appropriate group of experts and defend that choice.’
Table 1. Aid for researchers to explore the conceptualisation of an ‘expert’

1. What is your definition of an ‘expert’?

2. What type of Delphi is being utilised and what effect has this exerted on choice of expert?

3. What sample are you aiming for (homogenous or heterogeneous)?

4. How has the sampling method influenced your choice of experts (snowballing etc)?

5. What are your inclusion criteria, with justification for inclusion (Walker and Selfe 1996)?

6. What are your exclusion criteria, with justification for exclusion (Walker and Selfe 1996)?

7. How do you define knowledge? What level is required and how can this be identified?

8. How do you define experience? What level is required and how can this be identified? If experience has been defined through x number of years, is this defendable?

9. How do service users/carer/patients feature within the study? If excluded, why and how will their views be taken into account?

10. Were non-participants followed up? (Mullen 2003)

11. Within publications each expert needs to be clearly labelled. Walker et al (2000) defined their sample in terms such as ‘non-funding GP’ and ‘academic’ to enable expertness to be understood.

Table 1 shows an aid that is intended to help potential researchers to discuss, choose and, more importantly, defend their decisions for the selection or rejection of experts. Until clear consensus appears within the literature, researchers need to be able to justify their decisions in order for readers to ascertain the expertness of the panel. Current literature fails to defend the pros and cons of the expertness of a panel selection.

John Baker MPhil, MSc, B Nurs (Hons), RNM, is Health Foundation research fellow; Karina Lovell PhD, MSc, BA (Hons), RNM, is Professor; Neil Harris PhD, RNM, is nurse consultant/lecturer; all are based at the School of Nursing, Midwifery and Social Work, University of Manchester, UK.

This article has been subject to double-blind review


Online archive

For related articles and author guidelines visit our online archive at [www.nurseresearcher.co.uk](http://www.nurseresearcher.co.uk)