The Delphi Method: methodological issues arising from a study examining factors influencing the publication or non-Publication of mental health nursing research

ASHMORE, Russell, FLANAGAN, Tracy, MCINNES, Doug and BANKS, David

Available from Sheffield Hallam University Research Archive (SHURA) at:
http://shura.shu.ac.uk/12935/

This document is the author deposited version. You are advised to consult the publisher's version if you wish to cite from it.

Published version


Copyright and re-use policy

See http://shura.shu.ac.uk/information.html
The Delphi Method: methodological issues arising from a study examining factors influencing the publication or non-Publication of mental health nursing research
Abstract

Purpose – The paper describes how the classic Delphi method can be adapted and structured to ensure that specific research questions are clearly addressed.

Design/methodology/approach – As part of a larger mixed method project, a modified Delphi study was undertaken to explore factors influencing publication and non-publication of mental health nursing research.

Findings - This paper reports brief findings from the Delphi study. However, its main focus is the methodological issues arising from the Delphi method.

Implications - The paper argues that the classic Delphi method can be adapted and structured to ensure that specific research questions are able to be clearly answered. The adaptations are pragmatic in approach and in keeping with the general principles underpinning the Delphi method, while successfully addressing the problems of attrition and previous criticism of homogenous panels.

Originality/value - This paper offers some practical solutions to issue arising from undertaking research using the Delphi method.

Keywords: Barriers to publication, Modified Delphi method, Research, Mental health nursing.

Paper type: General review
Introduction
This article will discuss the Delphi methods employed in a study examining the reasons for mental health nursing research subsequently being published or not published. The article describes how the classic Delphi method can be adapted and structured to ensure that specific research questions are clearly addressed. The adaptations described in this article can be viewed as pragmatic in approach and in keeping with the general principles underpinning the Delphi method, while successfully addressing some of the problems commonly encountered when using this method. The principles of the classic Delphi method and how they were adapted to meet the needs of one research study will be discussed in this article.

Background
Only a small proportion of research carried out by nurses, the biggest health professional group within the field of mental health, is accessible through nursing publications (Hicks 1993, Hicks 1995). Yet there have been few studies exploring why nurses do not publish their research findings. In the most comprehensive group of studies carried out in the United Kingdom (UK), Hicks (1993) noted that only 0.4% - 3% of all nurses had ever published a research study. Hicks (1995) reported that 71% of her sample of nurses had carried out some research, 58% of this group had written up their work, and only 10% had submitted it for publication. Only 9% of those that had carried out research had their work published. Although it is difficult to judge the quality of any unpublished research, the likelihood is that a number of high quality mental health nursing research projects are unavailable to practitioners to aid their delivery of evidence based care.

Hicks’s (1993, 1995) work continues to be cited by authors (for example, Wilkes et al., 2002; Bishop, 2006; De Guzman et al., 2006; Stone et al., 2010), however it is now dated. A search of the literature failed to identify any recent research relating to this issue. However, the failure to publish research is a problem not confined to nursing. Several authors have reported on the number of research abstracts delivered at medical conferences that going on to be published in peer-reviewed journals. For example, Hackett et al. (2014) reported that only between 30.3% and 50.2% of abstracts presented at the International Liver Transplantation Society Conference between 2004 and 2008 were subsequently published in an academic journal.
In light of these concerns a project was established to examine those factors influencing the publication of UK based mental health nursing research. The project utilised a three-stage mixed methods design (Creswell and Plano Clark, 2011) consisting of a Delphi study, a survey, and semi-structured interviews. This article summarises the Delphi study and then examines the main methodological issues arising from the approach. The style adopted in this article, that is providing a summary of a study’s findings before reflecting on the techniques used to generate the data, is one that is commonly used in the literature (see for example, Gardner and Randall, 2012; Moore, 2012; Jessiman, 2013; Sangster-Gormley et al., 2015). Reflecting and reporting on issues encountered during field work has the potential to help future researchers to find practical solutions to their own methodological problems when using similar methods.

Summary of the Delphi study
The aim of the Delphi study was to identify what factors were perceived as important in contributing to the publication or non-publication of mental health nursing research. The methodology comprised a three-round modified Delphi method. The first round was undertaken during a workshop delivered at the International Network for Psychiatric Nursing Research (NPNR) conference. The workshop was also the means by which a purposive sample of ten experts was recruited. The panel was knowledgeable in undertaking and publishing mental health nursing research. During the workshop participants were each asked to identify up to five factors in answer to the question: ‘Why does only a small proportion of the mental health nursing research undertaken end up being published?’

Thirteen statements were derived by content analysis from participants’ responses. The statements, which reflected the potential reasons for why some research is not published, were returned to the experts for comment. Following slight adjustments the statements were sent out to the panel for scoring using a seven point Likert scale. Scores were collated and analysed to establish the mean and range for each statement. These findings were then sent back to the panel so that they could see their own scores and the overall scores of the panel. This offered them an opportunity to adjust their own scores if desired. The mean and range for each statement was then
recalculated. Statements were then separated into four groups related to their level of importance; high importance, importance, some importance and limited importance.

The statements considered of high importance when examining a lack of mental health nursing publications were:

- There is a lack of support when getting a paper ready for publication.

  “Support in terms of guidance from an academic supervisor or time off from work might well help people get published, but I have also seen people fall by the wayside even when this has been provided when they get asked to do the fifth re-write”.

- There is a lack of guidance on how to publish a paper.

  “General lessons or courses of publishing are giving only general guidance, more personal guidance is definitely needed to go through the process. Writing together with seniors could help. Workshops with guidance could be also one solution”.

One statement was considered of importance:

- The researcher lacks the motivation to pursue publication.

  “Unless your job and your career progression depend on it, publication is likely to slip from view sometimes as other areas of work come to assume great prominence”.

Statements considered of some importance by the panel were:

- The quality of the research is poor.
• Too much time has elapsed from the completion of the research to submit a publication.

• The research has been done for another purpose (such as an academic qualification or as a workplace task).

• An inexperienced researcher is unlikely to think about publication of their research.

• There is a lack of confidence in the standard of the research.

• The researcher chooses other ways of disseminating their research such as presenting at conferences.

• The publishing criteria are too stringent.

• An inappropriate journal is selected to publish the research.

Factors considered of limited importance were:

• There is no tradition of publication in mental health nursing.

• There is an academic/professional elitism that excludes some work from being published.

In summary the expert panellists saw lack of support and guidance as the most significant factors affecting research publication, lack of motivation to publish was also considered important. The factors generated in the Delphi study were incorporated into the semi-structured interviews undertaken in stage three of the study. They also provided the initial framework for undertaking the content analysis of these interviews.
The remainder of the article now focuses on the methodological issues raised by the Delphi study.

**The Delphi method**

The Delphi method has grown in popularity as an established research approach since its inception in 1963 by the RAND Corporation (Dalkey and Helmer, 1963). It is a useful method for inquiring into areas where there is little previous knowledge. It aims to explore and establish increased understanding through the process of seeking consensus on key issues from a group of identified experts. This is achieved through an iterative process; pertinent issues are identified and considered by the experts through successive rounds of questionnaires in which these issues are prioritised and ranked until consensus is achieved. The approach has evolved over time and the classical Delphi method (Table 1) has been modified and adjusted (Keeney, 2010; Hasson and Keeney, 2011), becoming an increasingly popular approach within health and social science and latterly nursing research (Beech, 2001; Powell, 2003; Baker et al., 2007; Parente and Anderson-Parente, 2011; Clibbens et al., 2012).

The approach allows greater freedom for experts to express their views and opinions freely, by ensuring that their responses are kept at least initially anonymous from other experts. This is achieved by corresponding with experts individually and then sharing their anonymous responses with the rest of the panel members. This mechanism of controlled feedback allows panellists to adjust their opinions by comparing their own positions against the rest of the experts. It thus avoids problems associated with other consensus methods such as focus groups or nominal group methods where dominant contributors may inhibit or contradict individual views (Keeney et al., 2001; Powell, 2003).

The Delphi method has been criticised for lack of validity, reliability and empirical rigour (Sackman, 1975; Williams and Webb, 1994; Keeney et al., 2001; Powell, 2003). Concerns have been raised about key elements of the process including; the definition and selection of experts (Baker et al., 2006), the means by which consensus is achieved (Greatorex and Dexter, 2000), and the objectivity of responses being undermined by potential failure to retain anonymity for expert panellists (Williams and Webb, 1994). It has also been argued that movement towards consensus by the
panellists may be governed by conformist and normative influences rather than an objective consideration of new information received and considered through the process of rounds (Goodman, 1987). These issues will now be considered in relation to the Delphi study summarised above.

**Defining and selecting the experts**

Central to the Delphi method is the identification of experts who can provide access to information on the issue under investigation (see Table 1). There is much debate in the literature as to what constitutes an expert; see for example Keeney (2010). Expertise may be conferred through formal professional or academic qualification. However, this negates expertise developed through experience, a particularly pertinent point to consider in the current climate of service user engagement and involvement with mental health research (Baker et al., 2006).

Attempts to clarify the definition of experts in the context of the Delphi approach has led to alternative terms including; ‘informed individual’, ‘specialist in the field’, ‘someone who has specialist knowledge about a specific subject’, and ‘informed advocates’ (Crisp et al., 1999; Keeney et al., 2001; Mead and Mosely, 2001; Baker et al., 2006). It seems essential that whichever definition of expert is accepted the person or persons need to be recognised as such, both within their peer or professional group and to a wider audience if the outcomes are to be deemed credible.

Some authors have suggested that the expert panel should be randomly sampled and representative in order to demonstrate generalisability of findings (Williams and Webb, 1994) or that a larger panel is favourable to establish reliability and validity (Murphy et al., 1998). This view is contested by Powell (2003) who argued that representativeness needs to be considered, not from a statistical perspective but from the relevant qualities and experiences of the panel members. Similarly, panel size should be determined by the nature of the subject of inquiry and availability of expertise.

Williams and Webb (1994, p. 182) found a paucity of studies stating any sort of inclusion criteria for expert panellist; they commented:
‘The notion that panel members are experts seem to be implicit in the fact that they are singled out for inclusion rather than fulfilling any specific standards’.

Hasson et al. (2000, p. 1008) cited purposive sampling or criterion sampling of experts as a preferred approach in many Delphi studies; although representativeness may be sacrificed as the panel is ‘selected for a purpose, to apply their knowledge to a certain problem on the basis of criteria’.

The focus of our Delphi study meant that only a relatively small group of individuals could be considered as experts. These included mental health nurses who had undertaken research, those who had subsequently gone on to publish, and those involved in reviewing research for publication. A leading mental health nursing research conference was identified as a useful site to recruit potential panellists. Therefore, it was decided to facilitate a workshop for delegates at the annual NPNR conference. The aims of the workshop were to introduce potential participants to the Delphi approach and to ask them to provide their opinions on the following question; ‘Why does only a small proportion of the mental health nursing research undertaken end up being published?’ Participants were then asked if they would be interested in becoming expert panellists, those expressing an interest were subsequently selected if they met the following criteria:

- They were mental health nurses that had undertaken research;
- The research approach/methodology utilised in their studies could be clearly defined and recognised as research (compared with, for example, clinical audit);
- They had experience of presenting and/or publishing research.

The workshop therefore provided opportunities to recruit experts to the project and to complete the first round of the Delphi study. Participants were also asked to suggest other peers or colleagues who might be interested in becoming expert panellists; this generated one further panel member who had not attended the workshop. This
The snowball method of sampling has been legitimised and used in other Delphi studies (see for example Baker et al., 2007). In total 10 expert panellists were recruited from higher education institutions within the UK and Europe. This number is consistent with the recommended optimum panel size (Baker et al., 2007). The workshop also provided participants with an opportunity to gain an understanding of the project and the process involved in generating data. In addition, it allowed potential experts to meet the research team which McKenna (1994) suggests is a positive factor in reducing attrition rates in subsequent rounds.

**Heterogeneous versus homogeneous panels**

Our intention was that the expert panel would comprise people who had successfully publishing their research and those who had not attempted to get their work published. However, following recruitment our expert panel was comprised of individuals who had successfully published, alongside those who were reviewers and/or editors for journals.

It has been suggested that there is benefit and strength to be gained from the selection of a heterogeneous panel with diverse and varied perspectives (Boulkedid et al., 2011). However, Jones and Hunter (1995) suggested that homogeneity with regard to clinical or professional expertise could be beneficial if the focus of the study related to clinical intervention or practice. They argued that such homogeneity would not necessarily preclude divergence of views and opinions but would provide credibility through equivalent knowledge experience and training in a specific area.

For the purposes of this study it was felt that inclusion of non-published members on the expert panel might have further illuminated the research question. Nevertheless, the knowledge and experience of those who had published and who were involved in reviewing and selecting papers for publication enabled us to be clearer about those factors that obstructed people from successfully publishing. The intention of our Delphi study was to seek to consolidate and expand understanding of factors contributing to publication. Therefore, it was felt justifiable to precede with a more homogenous panel providing credible responses and opinions based on expertise in publication of their own and others research.
Anonymity
Maintaining the anonymity of the expert panel is described as a central tenet of the classic Delphi method by some commentators (Goodman, 1987; Williams and Webb, 1994). Anonymity it is argued allows for the elimination of subject bias as individuals are afforded an equal opportunity to contribute and respond to ideas without responding to the perceived social pressure of knowing who the other panellists are (Gnatzy et al., 2011). Therefore, a potential weakness of our approach to recruitment was that it compromised anonymity of the future panel members as they were aware of who else had been in the workshop. In recognition of this participants were asked to; indicate their interest in becoming an expert panellist, give details of their research experience, and submit responses to the research question in writing without conferring with other participants in the workshop. Therefore, whilst they might have been aware of other potential panellists, ultimately they did not know who had put themselves forward or who had been selected. Similarly they did not necessarily know how other participants had responded.

However, Keeney et al. (2001) argue that anonymity is an unnecessary condition within a modified Delphi approach. Furthermore, Goodman (1987) suggested that complete anonymity might lead some panellists to make poorly considered or flippant contributions. Keeney et al. (2001) suggest that a level of cognisance between the panellists might actually increase their level of motivation to participate. Therefore, it is suggested that the key is to maintain anonymity of responses, so that experts are not aware of the provenance of other participants’ responses that they are asked to consider. As some panellists in our study were known to each other and in some instances worked together, complete anonymity would have been very difficult to achieve. However, panel members were requested not to discuss their responses with each other. Certainly the ability to complete rounds by electronic mail, which made the Delphi process easier and arguably reduced attrition rates in this study, would have been impossible if complete anonymity had been maintained.

The research community in mental health nursing is relatively small. This meant that several of the expert panellists were well known to the research team. This situation is strongly cautioned against by some (for example, Murphy et al., 1998) and this was initially acknowledged as an undesirable potential weakness. However, throughout
the subsequent rounds of the study the relationship of the research team to the panellists proved to have several benefits. The informality of the communication between rounds enabled tailored prompting and encouragement for panellists to return their responses. Knowledge of panellists’ workloads and other commitments allowed flexibility in the timescales given to respond. This meant that people were not arbitrarily excluded when they did not respond to reminders within a prescribed timescale.

Arguably the relationship of the panel to the research team might have compromised response objectivity. However, none of the panellists had a supervisory or managerial relationship with the team which mediated against any sense of obligation or coercion to remain in the study. The team’s relationship appeared to permit them to respond in a frank and honest manner, providing more relaxed and informal views alongside more considered and intellectually driven opinion.

**Attrition**

The Delphi method consists of a number of rounds in which questionnaires are sent to the panel until consensus is reached. During this process it is common for panel members to be lost; however the study outlined above had a zero rate of attrition, which is rare (Keeney, 2010). As noted above, this could be attributed to the personal nature of the relationship between the panel and the research team. This relationship seemed to make the process more enjoyable and less arduous. Other factors that could have resulted in zero attrition include; motivation to assist fellow researchers, a broader commitment to enquiry, and possibly a desire to develop experience of the methodology.

To minimise attrition it is also important to gauge the number of rounds people are asked to complete and the overall timescale of the study. From start to finish our study lasted nearly 12 months and comprised three rounds, including the first round undertaken in the conference workshop. Retaining participants also means anticipating amount of time, effort and engagement expected of them, efforts should be made to ensure that this is acknowledged and not exploited. Keeney et al. (2001) suggest that attrition can be minimised by maintaining the panel’s level of motivation, involvement and ownership of the subject under consideration. To this end,
participants were kept informed of the progress of the analysis and findings were disseminated as quickly as possible in an attempt to reflect our respect and gratitude for the contribution they have made (Gnatzy et al., 2011).

Consensus

The aim of the Delphi method is to gain consensus among a group of experts on a topic. However, there is much discussion in the literature about how this can be achieved and/or measured through Delphi techniques (Keeney, 2010). Within the classical Delphi method consensus usually reflects complete convergence within the rating of items (Linstone and Turoff, 1975). This would take four rounds to achieve and see the loss of more weakly scored items between rounds. There is an acknowledged risk that the greater the number of rounds the greater the risk of attrition from the expert panel especially in the final round. Williams and Webb (1994) stated that for consensus to be reached it is essential that all experts starting the process complete the final round. This then creates a double bind whereby zero attrition is a requirement for consensus but the number of rounds required to reach consensus increases the risk of attrition (Buck et al., 1993; Keeney et al., 2001). Modified versions of the Delphi method have introduced fewer rounds and/or adapted the definition of consensus to counter this double bind.

Williams and Webb (1994) reviewed different approaches to defining and determining of consensus. They found examples including the use of a predetermined level of consensus based upon a minimum percentage of respondent agreement and consideration of the measure of standard deviation for rated items. Responses with larger standard deviations were interpreted as having weaker consensus. These approaches allow for consensus to be determined and described in terms of those factors that are deemed as high and low importance to the panel.

Williams and Webb (1994) go on to suggest that in order to strengthen the reliability and validity of the findings there is a need to have a clear, predetermined definition of consensus to avoid agreement being reached through the researcher’s subjective view and manipulation of the data. However, Powell (2003) has noted that a number of studies have not attempted to demonstrate consensus through empirical means but have opted instead to present their findings and allow the audience to draw their own
conclusions about whether they are consensual or not. Powell (2003) suggested that striving for more ‘scientific’ legitimacy reflects an established trend towards deference to the dominant positivistic paradigm and, in addition, argued that robust and transparent decision making and execution of the approach might more usefully provide evidence of a study’s trustworthiness.

No predetermined level of consensus was set at the start of our study. Instead the mean and range of scores were calculated to enable identification of those items that were consistently scored highly; this also allowed those items that were scored more variably to be acknowledged through the rounds and in the final findings. A criticism of the Delphi approach is that items which achieve mixed or low scores (based on a calculation of the mean) may be excluded even though they may be factors contributing to a wider understanding of the subject under investigation (Hasson et al., 2000). However, this approach might be justifiable where large numbers of items for consideration may hinder the process of reaching consensus. However, given the narrow range of factors for consideration it was not considered necessary to exclude weaker items, as all of the items were felt to have some relevance. Therefore, the study’s findings were presented as those factors deemed to be of high importance, importance, some importance and limited importance (see above).

The process by which experts shift their view has also been explored with suggestions that normative as well as informative influences may be experienced. Therefore, experts may reconsider their view based on information they receive regarding the other panellists views. However, it is also acknowledged that they may adjust their position in light of the dominant view expressed by others and through a lack of confidence in their own opinion (Greatorex and Dexter, 2000). There is arguably only a subtle distinction between constructively reviewing one’s own opinion based on those of others and conforming to the rest of the group regardless of whether it is a view you would have reached autonomously. The reasons for adjusting or adopting a view may not even be consciously felt by the individual. Greatorex and Dexter (2000) acknowledge that those who refuse to adjust their position can be just as detrimental within a study as those who conform.
In our study some panellists shifted significantly in their view whilst others resolutely ‘held firm’. Some panel members reported that they had not expected to be influenced by the other panellists’ scores, but they then found themselves deciding to change their own scores after receiving feedback. There was insufficient feedback to gauge whether they were ‘constructively aligning’ or merely conforming. No feedback was elicited from those who had not changed their views across the rounds.

**Conclusions**

The decisions taken in planning and undertaking Delphi studies are infrequently reported in the literature. The challenges identified in undertaking the Delphi study described in this article meant that there was a need to modify the classic method (see Table 1) in order to ensure that specific research questions were clearly answered. Maintaining the anonymity of panel members is a central tenet of the Delphi method, however the nature of our study and the recruitment strategy employed meant that this could not be completely achieved. However, we have suggested that any potential compromise to objectivity this created was compensated by the researchers utilising a personalised approach for individual participants that ensured their continued engagement and allowed them to be frank and open. In addition, the zero attrition rates would seem to vindicate this approach.

In relation to the selection of the experts, the benefits afforded by having a more homogenous panel meant that a coherent view about reasons for non-publication could be achieved. This incorporated their personal experiences of trying to get published along with those factors that had supported and/or inhibited this process. Their expertise also provided a view from the perspective of reviewing and editing papers submitted for publication. The panellists’ responses provided invaluable insight into the kind of interventions that might be developed to enable more mental health nurses to publish their research. However, further understanding of these issues could have been enhanced by the recruitment of panellists who have not considered publishing their research or those who have tried and given up.

The utility of the Delphi approach has been demonstrated through the description of its application within a study exploring the reasons for poor publication rates of mental health nursing research. We believe that our experiences of careful and
considered decision making when using a modified Delphi design demonstrate the benefits of its flexibility and accessibility do not have to be at the cost of trustworthiness. Finally, it is suggested that the insights outlined in this article are of potential value to future researchers considering using the Delphi method.
References


Jessiman, W.C. (2013), “‘To be honest, I haven’t even thought about it’ – recruitment in small-scale, qualitative research in primary care”, Nurse Researcher Vol. 21, No. 2, pp. 18-23.


Table 1 Characteristics of the classic Delphi technique

1. A panel of ‘experts’ is used to obtain data.
2. Participants do not meet in face-to-face discussions.
3. The use of sequential questionnaires and/or interviews.
4. The emergence of a consensus of judgement/opinion.
5. The guarantee of anonymity for participants’ responses.
6. The quantification of earlier findings to identify patterns of agreement.
7. The use of two or more rounds between which the main findings of the previous round are communicated to and evaluated by panel members.

(See McKenna, 1994; Hsu and Sandford, 2007; Iqbal and Pidon-Young, 2009)