



Revisiting the Delphi technique - Research thinking and practice: A discussion paper

Felicity Hasson, Sinead Keeney^{*}, Hugh McKenna

School of Nursing and Paramedic Science, Ulster University, United Kingdom of Great Britain and Northern Ireland

ARTICLE INFO

Keywords:

Delphi technique
Methodology
Nursing
Consensus
Experts
Sampling

ABSTRACT

The Delphi technique is a research methodology which has traditionally been used to gain consensus among experts on complex issues characterised by uncertainty. Pioneered by the Rand Corporation in the 1950s for military applications, it has since been widely adopted across various fields, including nursing, health and social sciences and information systems on an array of multifaceted real-world issues. However, since its inception, the Delphi technique has undergone substantial methodological development and its use has now gone beyond its initial rationale.

In the last two decades there has been a growing body of work illustrating an increasing methodological diversity of the method. While such diversity presents possibilities, it also challenges traditional application and methodological rigour. In an attempt to preserve the integrity of the method, generic and discipline specific guidelines have emerged providing general principles and standards.

The aim of this paper is to present a much-needed critical reflection on the current application of the Delphi technique and its methodological development and to build on our paper from 2001 (Keeney et al., 2001). While the development of the Delphi method and its evolution are well recognised and reported in the literature, some controversies surrounding the approach remain and it is timely to revisit the method with a critical eye.

Ultimately, the Delphi technique's flexibility is its significant strength, enabling the exploration of novel lines of inquiry, but it also presents a challenge. Striking the right balance between flexibility and rigour can lead to more meaningful insights and actionable outcomes from a Delphi study. Yet to achieve this, some level of consensus may need to be reached on the Delphi technique itself. In recognition of its 60th birthday, it is an opportune time to re-examine its key aspects and methodological advances and reflect on 'when is a Delphi not a Delphi?'

What is already known

- The Delphi technique has been heralded by researchers in multiple disciplines as a means of achieving a convergence of opinion on an array of multifaceted real-world issues.
- The generic premise of a Delphi study is to generate, refine and synthesise opinions in a systematic manner using multiple structured rounds.
- The technique is an important and useful consensus methodology and has much to offer in gaining consensus from experts where none exists and it performs strongly in structuring group communication and decision making.

What this paper adds

- An update on our original critique paper from 24 years ago; this update reflects critically on the wide range of changes and variations in the type, use and application of the method that currently exists.
- An exploration of the substantial methodological development the Delphi technique has undergone in the past two and a half decades.
- A critical evaluation of the difficulties faced by researchers due to the wide array of amendments made to the Delphi technique, such as issues with consensus levels, expert panel selection, feedback mechanisms, and reliability.

^{*} Corresponding author.

E-mail address: s.keeney@ulster.ac.uk (S. Keeney).

<https://doi.org/10.1016/j.ijnurstu.2025.105119>

Received 4 February 2025; Received in revised form 2 May 2025; Accepted 9 May 2025

Available online 12 May 2025

0020-7489/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Since its creation in the 1940's, the Delphi technique has been heralded by researchers in multiple disciplines as a means of achieving a convergence of opinion on an array of multifaceted real-world issues. The term Delphi originates from Ancient Greece, where people sought guidance from the Delphic Oracle, conducted through Princess Pythia (a major oracle). This guidance was sought principally for forecasting the outcome of future wars or political actions. Whilst the original authors of the technique did not explain the origin of the name, in the 1950's the RAND Corporation developed the 'Delphi Method'. This is an exploratory social research technique which seeks to elicit group opinion using a structured and iterative approach (Dalkey and Helmer, 1963).

The technique was designed to achieve consensus on an issue by experts without direct confrontation. It reduced the risk of social pressure, which was considered a key shortcoming of other forecasting methods. Its initial application was part of the "Prediction Study" (later referred to as "Project Delphi") to help the United States Air Force formulate more reliable forecasts of future political developments.

Despite the method being ranked at the lowest level of evidence for making causal inferences (US Department of Health, 1992), its application and publication has escalated across diverse scientific fields. Since the first paper was published in defence and military matters (Dalkey and Helmer, 1963), estimates now suggest that the method has been utilised in 19,831 empirical English language studies with over 600 methodological papers being published, mainly in medical, science and technology and social science fields (Khodyakov et al., 2023). The popularity of the method has been attributed to its effectiveness in achieving a convergence of opinion among experts on complex, multifaceted issues, its adaptability allowing researchers to tailor it to various disciplines and contexts, making it applicable across diverse fields such as nursing, health sciences, social sciences and information systems and wide applicability in forecasting, evaluating knowledge, identifying priorities, developing guidelines, assessing policies and enhancing stakeholder engagement (Loë et al., 2016; Keeney et al., 2011; Donohue and Needham, 2008; Hasson et al., 2000; Keeney et al., 2001).

Over the last three decades the method has continued to play a pivotal role informing knowledge, understanding, and response across diverse scientific fields. Despite the increase in published guidelines and academic networks focusing on its development (e.g. DEWISS), many questions regarding this methodology continue to intrigue researchers. Given the increasing volume of literature on the method and in recognition of the Delphi technique's 60th anniversary, it seems an opportune time to re-examine its key aspects and methodological advances.

We published a well cited paper in this journal 24 years ago (citation number = 4075 Scopus). In the interim, we have watched its modifications and applications. This paper is an update on our original, exploring the wide range of variations in the type, use and application of the method. The Delphi technique remains an important and useful consensus methodology and has much to offer in gaining consensus from experts where none exists previously and it performs strongly in structuring group communication and decision making (Keeney et al., 2001). This paper is intended to critically evaluate the technique and offer insight into the methodological difficulties which need to be highlighted. It is not intended to be a guide to using the technique.

2. Delphi variability

Since its inception, the generic premise of a Delphi study has been to generate, refine and synthesise opinions in a systematic manner using multiple structured rounds. Traditionally, it has four key characteristic features, **anonymity, iteration with controlled feedback, statistical group response, and input**. These characteristics separate it from other group decision making processes (Rowe and Wright, 1999). The iterative nature of the process ensures that feedback is progressively

refined to achieve consensus or identify divergence. The original format of the Delphi is referred to as the 'Classical Delphi Technique', where the first round begins with an open-ended question asking individuals to identify issues of relevance to the topic, but where there is currently no agreement. These issues form the basis for subsequent rounds. Therefore, it is characterised by its iterative process designed to achieve consensus among a group of 'experts' through anonymous feedback and rounds. Its premise is that group opinion is more reliable than the opinion of individuals, whilst eliminating the negative effects of dominant personalities and group think (Dalal et al., 2011). Whilst the basic premise of the process remains, variations also exist that allow researchers the ability to address specific challenges, streamline the methodology, or align the approach with the study's objectives. Such variability is illustrated by Furtado et al. (2024) who assessed the level of variability of the Delphi used to gauge nursing competence. They identified 20 Delphi studies with diverse methodological options relating to the selection of experts, their input, analysis, consensus, and study procedures.

Originally the Delphi was administered via post, which was burdensome for both the researchers and respondents. Since then, the administration process has transformed with the use of technology, and specific online platforms host, collect and analyse Delphi results in real-time. Commonly referred to as e-Delphi (Keeney et al., 2011), they are justified by ensuring the inclusion of respondents across wide geographical areas and having greater convenience for panellists to participate. One of the first Delphi software products was developed in 1998, entitled Professional Delphi Scan (Kuusi et al., 2000). Since then, several real-time software-based platforms have emerged such as, Real-Time Delphi (RTD), Welphi, Mesydel, Calibrium, and eDelphi.org. In nursing, Bownes and Giannotti (2023) appraised critically the application of e-Delphi in postgraduate nursing education and competency development. They highlighted the variations in its application, but did not specify the types of administration used.

3. Delphi application

Since its inception the Delphi method has undergone substantial methodological development. The 'Classical Delphi' (Dalkey and Helmer, 1963) has transformed into multiple variations, departing widely from the original approach, causing no clear and uniform distinction in the literature on what a (genuine) Delphi variant is. Several authors have attempted to list the variations (Kendall, 1977; Hasson and Keeney, 2011; Mukherjee et al., 2015; Strasser, 2017). Seven variants were identified by Strasser (2017) including, Classic Delphi, Policy Delphi, Ranking-Type Delphi, Decision Delphi, Argument Delphi, estimate-feedback-talk-estimate (EFTE) Delphi, and Disaggregative Policy Delphi. Niederberger and Deckert (2022) has since listed an expansion of Delphi types to ten, including real time, markets, policy, argumentative, group, deliberative (citizens) Fizzy and Café, but note that some have rarely been used in practice.

Variations also exist with regard to how it is implemented such as Online Delphi, Real-Time Delphi, and Mini-Delphi (Hasson and Keeney, 2011; Strasser, 2017). More recently the technique has been blended with other approaches. For example, Khodyakov et al. (2024) have added an ExperLens that allows for implementing a modified Delphi online. It shares some features of Nominal Group Technique, enabling online discussion among participants that occurs between rounds.

Niederberger and Deckert (2022) advocated that variant election be aligned to the research aims, objectives and context of the study. However, its flexibility ensures that categories are not exhaustive or mutually exclusive and can be combined, depending upon the issue. Each variant may include some, or all the central features, of the Delphi (**anonymity, iteration with controlled feedback, statistical group response, and expert input**) but may differ in their goal, sampling, administration, analysis and outcomes. Adaptations of the Delphi have also been reported within variant types. For example, de Loë et al.

(2016) undertook a systematic review of 63 policy Delphi's published between 1971 and 2024 and reported wide variation in design, analysis and reporting techniques.

Guidance on how authors can adapt the technique exists (Brady, 2015), and variants therefore continue to evolve. A recent trend has been the use of the nested Delphi or the method used in combination with other approaches such as Q methodology (Wallis et al., 2009; Rust, 2017). Consequently, the true extent of the customisation of the Delphi and its process is unknown. Indeed, approaches that fall outside or in-between traditional thinking help to invigorate debates around what constitutes rigour and how rigour may be assessed outside of the current approaches. Yet the price of versatility, is heterogenous Delphi studies and methodological confusion, resulting in growing criticism for lacking scientific rigour, validity and transparency. Unsurprisingly, calls have been made for research to establish the conditions on which each Delphi variant is likely to be more effective (Bolger and Wright, 2011; Bolger et al., 2011; Rowe and Wright, 1999).

4. Defining the Delphi

At present, there is no universally accepted definition of the Delphi technique. Traditionally, it is commonly cited as a consensus-based method, applied when uncertainty exists. Consequently, initial definitions reflected the original intention of gaining consensus for future forecasting of real work practice and decision making. This is underpinned by the premise that anonymous group predictions are more valid than individual judgements. Seminal work by Dalkey and Helmer (1962, 1963) and Helmer (1967) defined it as obtaining a reliable consensus among a group of experts for judgmental forecasting of future events. Later Linstone and Turoff (1975) viewed it 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"* (p. 3). Since then, authors' operational definitions have been used interchangeably focusing on the goal of consensus making among experts (Hsu and Sandford, 2007; Jorm, 2015; Ab Latif et al., 2017) and describing the generic process of the method.

However, due to the universal characteristics of the method, its scope and application are much wider. Specific examples of its application range from: the evaluation of current knowledge; identification of convergence and dissonance of opinion; analysis of future developments; complementing or assessing feasibility studies; formulating theoretical or methodological guidelines, identification of priorities, identification of end points for clinical trials; assessment of the impact of policies, generation of solutions; development of assessment tools, standards, policies/guidelines and performance indicators and used as a means to enhance stakeholder engagement. It has also been used to enhance the rigour of other methods (Murphy et al., 2017; van Rijssen et al., 2019; Bull et al., 2022; Sakamoto, 2023).

As the Delphi method has evolved, an all-encompassing definition remains elusive. None exists that has kept abreast with the evolution of the process, the variants, differing administrations, and/or its application. Niederberger and Deckert (2022) have criticised narrow definitions focusing on the application as being untenable. However, the lack of nomenclature has the potential to interfere with the ability to compare findings across studies, identify standards for conducting Delphis and assessing and managing rigour.

5. Philosophical orientation

The Delphi technique is rooted in the philosophical assumptions of Locke, Kant and Hegel (Turoff, 1970). They emphasised the importance of the perceptions of a group of people, in conjunction with empirical data, in helping to establish reality. However, the philosophical paradigms underpinning the technique have been subject to some debate. This stems from the types of data the method can produce, bridging both the statistical analysis of quantitative data (aligned to positivism or post-

positivism) and qualitative statements (aligned to interpretivism or social constructivism) (Jaam et al., 2022; McIntyre-Hite, 2016). Originally, the technique was viewed through a pragmatic lens (Dalkey and Helmer, 1963) providing the researcher with the opportunity to use a range of strategies to answer the research questions. Being grounded in pragmatism, it underpins the key advantage of methodological flexibility and thinking, enabling the researcher to adopt the technique to the research context. Without guidance, the Delphi method has offered an opportunity for researchers to play within these boundaries, adapting approaches and developing variants of the technique that fit their epistemological stance, discipline, and specific research questions. Niederberger and Deckert (2022) believed that differing variants of the method have resulted in epistemological and methodological changes to the traditional understanding of the Delphi.

6. Delphi guidelines

In the last two decades there has been growing publication of methodological guidance for conducting Delphi studies. Building on the work of Spranger et al. (2022), Table 1 presents a summary of these reporting guidelines. Each guideline offers an insight into its structure, best practices, and areas where variations might occur depending on the research discipline. However, some report only on a certain component of the method or on the reporting of such studies. As Spranger et al. (2022) noted, none of these guidelines are applicable for the spectrum of Delphi variants or their applications in the different fields. Moreover, Banno et al. (2020) asserted that most Delphi guidelines published prior to 2019 were not developed with the Delphi methodology. Therefore, in the absence of generic all-encompassing guidance, such guidelines are used in conjunction with other methodological checklists such as the Checklist for Reporting Results of Internet E- Surveys Checklist (CHERRIES).

7. Characteristics of the Delphi

The inherent characteristics of the Delphi technique remain controversial, and, for context, these are discussed briefly below. A more in-depth discussion of the characteristics of the Delphi can be found at Hasson et al. (2000) and Keeney et al. (2011).

7.1. Anonymity

A key feature of the Delphi technique is its claim that participants are anonymous to each other but not to the researcher, helping to address a range of social pressures that affect group-based decisions. McKenna (1994) coined the term 'quasi-anonymity'. He suggested that the Delphi technique collects data from expert panel members who may know one another, but that their judgements and opinions can remain anonymous. However, it remains unclear as to what influence anonymity has on the process or the results.

7.2. Consensus levels

No universal agreement exists on the approach to establish consensus, the level of consensus to be applied both within and at the end of Delphi rounds and little understanding of how it is reached among expert panels (von der Gracht, 2012; Boulkedid et al., 2011; Diamond et al., 2014). Commonly, consensus is presented, a priori, as a threshold percentage to illustrate the level of agreement of responses. However, it can range from 50 to 97 % (Barrios et al., 2021). Such arbitrary approaches may go some way to explain why Delphi researchers often refrain from defining consensus (Diamond et al., 2014).

7.3. Expert panel

The Delphi was originally developed as an expert elicitation

Table 1
Overview of Delphi guidelines.

Guideline title	Discipline	Author
RAND Methodological guidance for conducting and crucially appraising Delphi panels	Generic	Khodyakov et al. (2024)
Accurate Consensus Reporting Documents (ACCORD)	Biomedicine	Gattrell et al. (2022, 2024)
DELPHISTAR Standards for reporting Delphi's use in health and social sciences	Health and social sciences	Spranger et al. (2022)
Guidelines on the Delphi method appropriateness	Generic	Nasa et al. (2021)
White paper: Using the Delphi technique to achieve consensus on bereavement care in palliative care in Europe	Bereavement and Palliative care	Keegan et al. (2021)
COS-STAR (Core Outcome Set-STAndards for Reporting), Guidelines on the reporting of the consensus method used as part of the study	Generic	Kirkham et al. (2016)
Guidance on multiple aspects of Delphi process	Clinical trials in rheumatic diseases	Humphrey-Murto et al. (2017)
Reporting of Delphi methods to achieve consensus on guidelines in Rare Diseases	Rare diseases	Resemann et al. (2018)
How to use guide when considering planning and publishing studies using consensus methods	Medical education	Humphrey-Murto et al. (2017)
Guidance on Conducting and REporting DELphi Studies (CREDES) 2017	Palliative care	Jünger et al. (2017)
Reporting E-Delphi Studies (REDS) in health research: development of a preferred items checklist 2016.	Health research	Guerreiro et al. (2018)
Reporting and reading template for Delphi studies	Primary care	Toma and Picioreanu (2016)
Guidelines for describing the use of consensus methods.	Generic	Waggoner et al. (2016)
Guidelines for capturing and reporting expert opinion	Pharmacoeconomics	Iglesias et al. (2016)
How to report consensus in Delphi Studies	Generic	Diamond et al. (2014)
Guidance on how the selection of health quality indicators should be reported in Delphi studies	Health services	Boulkedid et al. (2011)
Guidelines on the development of outcomes for presenting measurements in clinical trials	Clinical trails	Sinah et al. (2011)
Guidelines for the Delphi technique - outline of the preparation, action, and steps of Delphi studies	Nursing	Hasson et al. (2000)

methodology (Dalal et al., 2011), hence the term 'experts' was commonly used. However, keeping pace with the many variants of the method, the term 'panellist' is increasingly being used. Regardless of terminology, the identification and selection of the panel is considered a crucial aspect of any Delphi study. Nonetheless, there is no standard definition or criteria in existence to determine the eligibility, selection, number or retention of panellists. For example, panel sizes can vary from 10 to 1000, while recommendations vary from 10 to 25 participants (Vogel et al., 2019), 30–50 panellists (McGrath et al., 2020; Robba et al., 2020) to 60–80 in multistakeholder studies (Manyara et al., 2024). Optimal panel size varies substantially based on the intended purpose of the Delphi study. Whilst small samples call into question the rigour of the findings (Niederberger and Deckert (2022)), large samples may suffer from the law of diminishing returns (Yoshida et al., 2016). Consequently, panel numbers are commonly discrepant in published Delphi studies, often related to the study's objective, complexity and resources available (Nasa et al., 2021).

7.4. Feedback

There are still no agreed guidelines or consistency on how many rounds are required within a Delphi study or what is the best approach to provide feedback. This is despite its role in influencing individual responses and the achievement of consensus (Barrios et al., 2021; Turnbull et al., 2018; Bolger et al., 2011; Rowe and Wright, 1999). The number of rounds used can vary from two to seven (Khodyakov et al., 2023) dependent on the study's resources and methodological decisions. Confusingly, even a survey of experts using a single round has been labelled as a Delphi study (Antonio et al., 2020).

Commonly, feedback provided is either a statistical summary, using a mean, a median or the percentage of agreement or argumentative qualitative feedback, outlining reasons (Boulkedid et al., 2011). Clarity on what approach should be adopted or the impact of feedback on individual opinions and consensus are unknown.

7.5. Statements and scales

Traditionally in a Delphi study, consensus is based on a series of statements, but despite their importance, little attention has been given to their development or measurement. In a classical design, the first round is qualitative consisting of open-ended questions (Hsu and Sandford, 2007; Skinner et al., 2015) that are transformed into statements upon which to collect levels of agreement in later rounds. In modified Delphis, pre-selected items, devised from various sources such as evidence synthesis, exploratory workshops, policies, guidelines and/or questionnaires are used to inform first-round Delphi statements. Whilst Anderson (2022) set out recommendations regarding the word length of statements, there exists a lack of practical understanding of the process. Indeed, responses to Delphi statements are normally framed as Likert scales, but there is no agreement on the number of scale response categories. Rather, scale response options range from 5-point (Smith et al., 2016; Dy et al., 2013), 7-point (Weng, 2004) to 9-point (Diamond et al., 2014), with little agreement on the use of neutral or midpoint option (von der Gracht, 2012). As such, the use of different scales raises questions regarding the reliability of the findings (Lange et al., 2020).

8. Reliability and validity

The Delphi technique has been persistently plagued by questions regarding its integrity. Several challenges and points of contention exist, that affect perceptions of the method. These centre on the absence of standardised rigour approaches in its testing and application. The method lacks a universally accepted set of rules or rigorous testing procedures. Belton et al. (2019) argued that it is wrong to continually apply the concepts of reliability and validity that have their origins in psychometrics, since the Delphi is not a psychometric instrument. Instead, it is an inherently versatile and subjective method, reliant on the opinions, judgements, and expertise of individuals. Unsurprisingly, challenges have been reported when assessing the reliability of the method. Traditionally, reliability has been evaluated by comparing two or more Delphi studies on the same subject, but evidence underpinning inter-observer reliability varies with some empirical support evident (Dagenais, 1978; Duffield, 1993; Kastein et al., 1993; Quintana et al., 2000). Some authors believe there is no way of confirming whether two or more panels who go through the same Delphi process would provide the same results (Hasson et al., 2000; Keeney et al., 2001). The stability or replicability of the Delphi has also been criticised (Shekelle et al., 1998), with some asserting that this is not an appropriate measure, as the underlying essence of a Delphi is to allow participants to change their opinion (Okoli and Pawlowski, 2004).

The lack of standard application leads to the level of rigour of individual studies being heavily dependent on the meticulousness of the design and execution of the study. This is dependent on a range of factors, including the selection of experts, the clarity of the questions posed,

the number of rounds conducted, the iteration process and the methodology employed to analyse the results. Several internal strategies are advocated to enhance reliability within individual studies such as careful consideration of panel selection, questionnaire development, round structure, format of feedback, facilitation, consensus thresholds, statistical analysis, engagement, and transparency to enhance robust reporting.

Whilst some authors suggest that content validity may be assured by the careful selection of the expert panel, with the composition being as representative as possible (Hasson et al., 2000), such approaches are hampered if high attrition between rounds occurs. Whilst Banerjee et al. (2020) reported that the value of a single reliability measure is limited, many studies continue to report stability in consensus levels and group stability in responses, through the calculation of the interquartile range (IQR), variation in the coefficient of variation (VCV) (rank order and intraclass), and kappa calculations (de Vet et al., 2013; Kastein et al., 1993). However, the appropriate level of agreement of kappa values is context specific (Kottner et al., 2011). Lange et al. (2020) maintained that such established psychometric values may be considered as a rule of thumb and used with caution.

There is no doubt that presently the Delphi technique has methodological difficulties, and this is being highlighted in this paper. However, it is important to note that all research approaches have generated criticisms about their inconsistencies and biases including randomised controlled trials (Fernainy et al., 2024), questionnaires (Bee and Murdoch-Eaton, 2016) and interviews (Alsaawi, 2014). It is important that research methods are viewed with a critical eye and in many cases, this can serve to improve the method by addressing these limitations. The definitive research method has yet to be invented. It is important to acknowledge also that as a means of achieving consensus on important issues, where none existed previously, the Delphi is one of the best approaches available.

There are an almost unlimited number of important issues in nursing and healthcare generally where agreement is lacking. Despite its many variations and methodological challenges, the Delphi Technique as a consensus methodology has many key strengths which make it a good choice for studies seeking expert opinion and consensus on issues and topics where none exists. The key foundations of the Delphi in its original form are testament to its longevity and its strengths still include: anonymity of experts ensuring unbiased responses by reducing dominance by any one individual and by reducing group conformity; an iterative process allowing expert panel members to refine their opinions in the context of other expert opinion which in turn improves the quality and reliability of consensus; its method of controlled feedback providing summarised data on an individual and group basis enabling informed adjustments to previous responses; the ability to build consensus in facilitating the convergence of expert opinion, particularly where evidence is limited; flexibility in applying itself to many wide and varied disciplines; and the ability to enable global participation on subjects using experts worldwide through use of technology.

9. Implications for nursing

As a relative newcomer into academia and research, there remains a lack of consensus on many of the phenomena that are important to nurses and nursing. The Delphi technique has emerged as a pivotal methodology for the discipline, facilitating the gathering of expert consensus on various issues within the field. Numerous studies in nursing have utilised the method; recent examples include Levett-Jones et al. (2024) who used the Delphi to achieving consensus on the essential knowledge and skills needed by nursing students to promote planetary health and sustainable healthcare; Varndell et al. (2021) employed the Delphi to generate guidance in emergency nursing practice; Paz-Pascual et al. (2019) did so to gain consensus on priorities in maternal education; and Sim et al. (2017) to measure the outcomes of nursing practice.

The structured approach has allowed nurse researchers to solicit

opinions from panels of experts through multiple rounds of questionnaires, ultimately refining responses and achieving a collective agreement. The anonymity afforded by the method can facilitate honest and uninhibited feedback, which can be particularly beneficial in sensitive areas of nursing policy and practice. However, while accepting its popularity, it is essential to take cognisance of the methodological challenges discussed above.

10. Conclusions

Since its inception the Delphi Technique has become a popular research method across disciplines including nursing, but it has been challenged methodologically. Advances have been made regarding the development of guidelines, its flexibility, and whilst a strength, is also its potential disadvantage. Despite its historical significance and adaptability over time, the technique's efficacy is contingent upon the expertise and engagement of its panel participants. Previous critical reviews have highlighted its methodological shortcomings arising through the many deviations from the original process (Beretta, 1996; Keeney et al., 2001; Donohue and Needham, 2008).

Since its development in the 1950s, several variations have been developed and continue to be developed, resulting in the classical elements upon which the method was based becoming increasingly unrecognisable. For example, a central tenant for the introduction of the method was anonymity, lauded for helping to remove the potential of bias due to group psychosocial pressure, however in some variants being used today, anonymity may not be present.

This paper has discussed the wide range of variations in the type, use and application of the method. The lack of standardisation, in definitions, processes and reporting approaches have been criticised, questioning the level of rigour or value of decisions made. Whilst many variants are listed, it is unclear if, and how, they are being used in practice. Since the Delphi's inception, it has attracted much interest from a methodological perspective, however the track record of its progression is mixed. While it has gained traction in nursing and health sciences research due to its ability to gain consensus from diverse perspectives, it remains constrained by inherent limitations. The quality of outcomes can vary significantly based on expert selection and adherence to methodological rigour. Although it facilitates consensus-building on complex issues with uncertain knowledge bases, researchers must be cautious in interpreting results as definitive evidence.

Significant progress has been made in the development of Reporting Guidelines for the Delphi Technique as described in Table 1 but nothing similar has been attempted to date in relation to criteria for the Delphi as a method. For example, are there certain characteristics that must be present for a Delphi study to be considered a Delphi study? Does quasi-anonymity have to be present alongside a certain number of rounds with group and/or individual feedback, an expert panel and a statistically stable level of consensus for it be considered a Delphi study? Could there be a set of defined criteria agreed upon for the definition of a Delphi or has the method become so widely varied that that would not be possible?

Delphi can, and has, provided valuable insights, and fostered informed decision-making in research in many fields including nursing. However, over the years, its versatility and subjectivity has brought a diversified application, leading to the quest for broader definitions, descriptions, and classifications. Ultimately, the Delphi technique's flexibility is its significant strength, but it is also a challenge. Striking the right balance between flexibility and rigour can lead to more meaningful insights and actionable outcomes from a Delphi study. To achieve this, we may need to pursue some level of consensus on the Delphi technique itself.

CRedit authorship contribution statement

Felicity Hasson: Writing – review & editing, Writing – original draft,

Visualization, Supervision, Resources, Project administration, Methodology, Conceptualization. **Sinead Keeney**: Writing – review & editing, Writing – original draft, Visualization, Supervision, Project administration, Methodology, Conceptualization. **Hugh McKenna**: Writing – review & editing, Writing – original draft, Visualization, Conceptualization.

Funding

This work did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

None.

References

- Ab Latif, R., Dahlan, A., Ab, Mulud Z., Mat Nor, M.Z., 2017. The Delphi technique as a method to obtain consensus in health care education research. *Education in Medicine Journal*. 9 (3), 89–102.
- Alsaawi, A., 2014. A critical review of qualitative interviews. *Eur. J. Bus. Soc. Sci.* 3, 149–156. <https://doi.org/10.2139/ssrn.2819536>.
- Anderson, P.D., 2022. Constructing Delphi statements for technology foresight. *Futures Foresight Sci.* 5 (2). <https://doi.org/10.1002/ffo2.144>.
- Antonio, C.A.T., Bermudez, A.N.C., Cochon, K.L., Reyes, M.S.G.L., Torres, C.D.H., Liao, S.A.S.P., 2020. Recommendations for intersectoral collaboration for the prevention and control of vector-borne diseases: results from a modified delphi process. *J. Infect. Dis.* 222, S726–S731.
- Banerjee, A., Splieth, C., Breschi, L., Fontana, M., Paris, S., Burrow, M., Crombie, F., Foster Page, L., Gatón-Hernández, P., Giacaman, R.A., Gugni, N., Hickel, R., Jordan, R.A., Leal, S., Lo, E., Tassery, H., Thomson, W.M., Manton, D.J., Schwendicke, F., 2020. When to intervene in the caries process? A Delphi consensus statement. *Br. Dent. J.* 229 (7), 474–482.
- Banno, M., Tsujimoto, Y., Kataoka, Y., 2020. The majority of reporting guidelines are not developed with the Delphi method: a systematic review of reporting guidelines. *J. Clin. Epidemiol.* 124, 50–57 (Aug).
- Barrios, M., Guirera, G., Nuño, L., Gómez-Benito, J., 2021. Consensus in the delphi method: what makes a decision change? *Technol. Forecast. Soc. Change* 163, 120–484.
- Bee, D., Murdoch-Eaton, D., 2016. Questionnaire design: the good, the bad and the pitfalls. *Archives of Disease in Childhood - Education & Practice Edition*. 101, edpract-2015. <https://doi.org/10.1136/archdischild-2015-309450>.
- Belton, I., MacDonald, A., Wright, G., Hamlin, I., 2019. Improving the practical application of the Delphi method in group-based judgment: a six-step prescription for a well-founded and defensible process. *Technol. Forecast. Soc. Chang.* 147, 72–82.
- Beretta, R., 1996. A critical review of the Delphi technique. *Nurs. Res.* 3 (4), 79–89.
- Bolger, F., Wright, G., 2011. Improving the Delphi process: lessons from social psychological research. *Technol. Forecast. Soc. Chang.* 78 (9), 1500–1513.
- Bolger, F., Stranieri, A., Wright, G., Yearwood, J., 2011. Does the Delphi process lead to increased accuracy in group-based judgmental forecasts or does it simply induce consensus amongst judgmental forecasters? *Technol. Forecast. Soc. Chang.* 78 (9), 1671–1680.
- Boulkedir, R., Abdoul, H., Loustau, M., Sibony, O., Alberti, C., 2011. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One* 6 (6), e20476.
- Bownes, N.A., Giannotti, N., 2023. E-delphi technique in postgraduate registered nursing education and competency development: a scoping review. *Quality Advancement in Nursing Education* 9 (1), 3.
- Brady, S.R., 2015. Utilizing and adapting the delphi method for use in qualitative research. *Int J Qual Methods* 14 (5), 1609406915621381. <https://doi.org/10.1177/1609406915621381>.
- Bull, C., Crilly, J., Latimer, S., Gillespie, B.M., 2022. Establishing the content validity of a new emergency department patient-reported experience measure (ED PREM): a Delphi study. *BMC Emerg. Med.* 22 (1), 65.
- Dagenais, F., 1978. The reliability and convergence of the Delphi technique. *J. Gen. Psychol.* 98 (2), 307–308.
- Dalal, S., Khodyakov, D., Srinivasan, R., Straus, S., Adams, J., 2011. ExpertLens: a system for eliciting opinions from a large pool of non-collocated experts with diverse knowledge. *Technol. Forecast. Soc. Chang.* 78, 1426–1444.
- Dalkey, N., Helmer, O., 1962. An experimental application of the delphi method to the use of experts. *Memorandum RM-727/1*-abridged. *Manag. Sci.* 9 (3), 458–467.
- Dalkey, N., Helmer, O., 1963. Delphi Technique: characteristics and sequence model to the use of experts. *Manag. Sci.* 9 (3), 458–467.
- de Vet, R., van Lijstelaar, M.J., Brilleslijper-Kater, S.N., Vanderplasschen, W., Beijersbergen, M.D., Wolf, J.R., 2013. Effectiveness of case management for homeless persons: a systematic review. *Am. J. Public Health* 103 (10), e13–e26.
- Diamond, I.R., Grant, R.C., Feldman, B.M., Pencharz, P.B., Ling, S.C., Moore, A.M., Wales, P.W., 2014. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J. Clin. Epidemiol.* 67 (4), 401–409.
- Donohue, H.M., Needham, R.D., 2008. Moving best practice forward: delphi characteristics, advantages, potential problems, and solutions. *International Journal of Tourism* 11 (5), 415–437.
- Duffield, C., 1993. The Delphi technique: a comparison of results obtained using two expert panels. *Int. J. Nurs. Stud.* 30 (3), 227–237.
- Dy, C.J., Gonzalez Della Valle, A., York, S., Rodriguez, J.A., Sculco, T.P., Ghomrawi, H.M., 2013. Variations in surgeons' recovery expectations for patients undergoing total joint arthroplasty: a survey of the AAHKS membership. *J. Arthroplast.* 28 (3), 401–405.
- Fernainy, P., Cohen, A.A., Murray, E., et al., 2024. Rethinking the pros and cons of randomized controlled trials and observational studies in the era of big data and advanced methods: a panel discussion. *BMC Proc.* 18 (Suppl. 2), 1.
- Furtado, L., Coelho, F., Pina, S., Ganito, C., Araújo, B., Ferrito, C., 2024. Delphi technique on nursing competence studies: a scoping review. *Healthcare (Basel)* 12 (17), 1757. <https://doi.org/10.3390/healthcare12171757>. PMID: 39273781; PMCID: PMC11395531. Sep 3.
- Gattrell, W.T., Hungin, A.P., Price, A., Winchester, C.C., Tovey, D., Hughes, E.L., 2022. ACCORD guideline for reporting consensus-based methods in biomedical research and clinical practice: a study protocol. *Research Integration Peer Review* 7 (3) (pmid:35672782).
- Gattrell, W.T., Logullo, P., van Zuuren, E.J., Price, A., Hughes, E.L., Blazey, P., 2024. ACCORD (ACcurate Consensus Reporting Document): a reporting guideline for consensus methods in biomedicine developed via a modified Delphi. *PLoS Med.* 2024 (21), e1004326.
- Guerreiro, M.P., Plácido, M., Barros, C.T., 2018. A national e-delphi towards the measurement of safe medication practices in Portuguese hospitals. *Eur. J. Hosp. Pharm.* 25 (2), 103–106. <https://doi.org/10.1136/ejpharm-2016-000955>.
- Hasson, F., Keeney, S., 2011. Enhancing rigour in the delphi technique research. *Technol. Forecast. Soc. Change* 78 (9), 1695–1704. <https://doi.org/10.1016/j.techfore.2011.04.005>.
- Hasson, F., Keeney, S., McKenna, H., 2000. Research guidelines for the Delphi survey technique. *J. Adv. Nurs.* 32, 1008–1015.
- Helmer, O., 1967. Analysis of the Future: The Delphi Method. Rand, Santa Monica, CA.
- Hsu, C., Sandford, B.A., 2007. The Delphi technique: making sense of consensus. *Pract. Assess. Res. Eval.* 12 (1), 10. <https://doi.org/10.7275/pdz9-th90>.
- Humphrey-Murto, S., Varpio, L., Wood, T.J., Gonsalves, C., Uffholz, L.A., Mascioli, K., Wang, C., Foth, T., 2017. The use of the Delphi and other consensus group methods in medical education research: a review. *Acad. Med.* 92 (10), 1491–1498.
- Iglesias, C.P., Thompson, A., Rogowski, W.H., Payne, K., 2016. Reporting guidelines for the use of expert judgement in model-based economic evaluations. *Pharmacoeconomics* 34 (11), 1161–1172.
- Jaam, M., Awaisi, A., El-Awaisi, A., Stewart, D. El, Hajj, M.S., 2022. Use of the delphi technique in pharmacy practice research. *Res. Soc. Adm. Pharm.* 18 (1), 2237–2248.
- Jorm, A.F., 2015. Using the delphi expert consensus method in mental health research. *Aust. N. Z. J. Psychiatry* 49 (10), 887–897. <https://doi.org/10.1177/0004867415600891>.
- Jünger, S., Payne, S.A., Brine, J., Radbruch, L., Brearley, S.G., 2017. Guidance on Conducting and REporting DELphi Studies (CREDES) in palliative care: recommendations based on a methodological systematic review. *Palliat. Med.* 31, 684–706.
- Kastein, M.R., Jacobs, M., van der Hell, R.H., Luttik, K., Touw-Otten, F., 1993. Delphi, the issue of reliability: a qualitative delphi study. *Technol. Forecast. Soc. Change* 44, 315–323.
- Keegan, O., Murphy, I., Benkel, I., Limonero, J.T., Relf, M., Guldin, M.-B., 2021. Using the Delphi technique to achieve consensus on bereavement care in palliative care in Europe: an EAPC White Paper. *Palliat. Med.* 35 (10), 1908–1922.
- Keeney, S., Hasson, F., McKenna, H.P., 2001. A critical review of the delphi technique as a research methodology for nursing. *Int. J. Nurs. Stud.* 38 (2), 195–200. [https://doi.org/10.1016/s0020-7489\(00\)00044-4](https://doi.org/10.1016/s0020-7489(00)00044-4).
- Keeney, S., Hasson, F., McKenna, H.P., 2011. The Delphi Technique in Nursing and Health. Wiley, London.
- Kendall, J.W., 1977. Variations of Delphi. *Technol. Forecast. Soc. Chang.* 11 (1), 75–85.
- Khodyakov, D., Grant, S., Kroger, J., Gadowah-Meaden, C., Motala, A., Larkin, J., 2023. Disciplinary trends in the use of the Delphi method: a bibliometric analysis. *PLoS One* 18, e0289009.
- Khodyakov, D., Grant, S., Kriger, J., Bauman, M., 2024. RAND Methodological Guidance for Conducting and Crucially Appraising Delphi Panels. RAND Corporation, Santa Monica, Calif.
- Kirkham, J.J., Davis, K., Altman, D.G., Blazeby, J.M., Clarke, M., Tunis, S., Williamson, P. R., 2016. Core outcomes set=STAndards for development: the cos-stad recommendations. *PLoS Med.* 14 (11), e1002447.
- Kottner, J., Audigé, L., Brorson, S., Donner, A., Gajewski, B.J., Hróbjartsson, A., Roberts, C., Shoukri, M., Streiner, D.L., 2011. Guidelines for Reporting Reliability and Agreement Studies (GRRAS). *J. Clin. Epidemiol.* 64 (1), 96–106.
- Kuusi, O., Hiltunen, E., Linturi, H., 2000. Weak signals – a Delphi study. *Futura* 2000 (2), 78–92.
- Lange, T., Kopkow, C., Lütznier, J., 2020. Comparison of different rating scales for the use in Delphi studies: different scales lead to different consensus and show different test-retest reliability. *BMC Med. Res. Methodol.* 20, 28. <https://doi.org/10.1186/s12874-020-0912-8>.
- Levet-Jones, T., Catling, C., Cheer, S., Fields, L., Foster, A., Maguire, J., McIntyre, E., Moroney, T., Pich, J., Pitt, J., Whiteing, N., Lokmic-Tomkins, Z., 2024. Achieving consensus on the essential knowledge and skills needed by nursing students to

- promote planetary health and sustainable healthcare: a Delphi study. *J. Adv. Nurs.* <https://doi.org/10.1111/jan.16229>.
- Linstone, H.A., Turoff, M., 1975. *The Delphi Method: Techniques and Applications*. Addison-Wesley, Reading, MA.
- Loë, R., Melnychuk, N., Murray, D., Plummer, R., 2016. Advancing the state of policy Delphi practice: a systematic review evaluating methodological evolution, innovation, and opportunities. *Technol. Forecast. Soc. Chang.* 104, 78–88.
- Manyara, A., Purvis, A., Oriana, C., Collins, G., Taylor, R., 2024. Sample size in multistakeholder Delphi surveys: at what minimum sample size do replicability of results stabilise? *J. Clin. Epidemiol.* 174, 111485. <https://doi.org/10.1016/j.jclinepi.2024.111485>.
- McGrath, B.A., Brenner, M.J., Warrillow, S.J., Pandian, V., Arora, A., Cameron, T.S., Añon, J.M., Hernández Martínez, G., Truog, R.D., Block, S.D., Lui, G.C.Y., McDonald, C., Rassekh, C.H., Atkins, J., Qiang, L., Vergez, S., Dulgerov, P., Zenk, J., Antonelli, M., Pelosi, P., Walsh, B.K., Ward, E., Shang, Y., Gasparini, S., Donati, A., Singer, M., Openshaw, P.J.M., Tolley, N., Markel, H., Feller-Kopman, D.J., 2020. Tracheostomy in the COVID-19 era: global and multidisciplinary guidance. *Lancet Respir. Med.* 8, 717–725. [https://doi.org/10.1016/S2213-2600\(20\)30230-7](https://doi.org/10.1016/S2213-2600(20)30230-7).
- McIntyre-Hite, L., 2016. A delphi student of effectiveness practices for developing competency-based learning models in higher education. *J. Competency Based Educ.* 1 (4), 157–166.
- McKenna, H.P., 1994. The Delphi Technique: a worthwhile research approach for nursing? *J. Adv. Nurs.* 19 (6), 1221–1225.
- Mukherjee, N., Hugé, J., Sutherland, W.J., McNeill, J., Van Opstal, M., Dahdouh-Guebas, F., Koedam, N., 2015. The delphi technique in ecology and biological conservation: applications and guidelines. *Methods Ecol. Evol.* 6 (9), 1097–1109. <https://doi.org/10.1111/2041-210X.12387>.
- Murphy, M., Hollinghurst, S., Salisbury, C., 2017. Agreeing the content of a patient-reported outcome measure for primary care: a Delphi consensus study. *Health Expect.* 20 (2), 335–348.
- Nasa, P., Jain, R., Juneja, D., 2021. Delphi methodology in healthcare research: how to decide its appropriateness. *World Journal of Methodology* 11 (4), 116–129.
- Niederberger, M., Deckert, S., 2022. The Delphi Method: Methodology, Variants and Application Examples Evidence in Health Care, 174, pp. 11–19.
- Okoli, C., Pawlowski, S.D., 2004. The delphi method as a research tool: an example, design considerations and applications. *Inf. Manag.* 42 (1), 15–29. <https://doi.org/10.1016/j.im.2003.11.002>.
- Paz-Pascual, C., Artieta-Pinedo, I., Grandes, G., 2019. Consensus on priorities in maternal education: results of Delphi and nominal group technique approaches. *BMC Pregnancy Childbirth* 19, 264. <https://doi.org/10.1186/s12884-019-2382-8>.
- Quintana, J.M., Arostegui, I., Azkarate, J., Goenaga, J.I., Elexpa, X., Letona, J., Arcelay, A., 2000. Evaluation of explicit criteria for total hip joint replacement. *J. Clin. Epidemiol.* 53 (12), 1200–1208.
- Resemann, H.K., Clements, S.J., Griffiths, A., Sheard, D., 2018. Reporting of Delphi Methods to Achieve Consensus on Guidelines in Rare Diseases Presented at the 2018 European Meeting of the International Society for Medical Publication Professionals (ISMPP), 23–24 January 2018, London, UK..
- Robba, C., Poole, D., Citerio, G., Taccone, F.S., Rasulo, F.A., 2020. Consensus on brain ultrasonography in critical care group. Brain ultrasonography consensus on skill recommendations and competence levels within the critical care setting. *Neurocrit. Care.* 32 (2), 502–511.
- Rowe, G., Wright, G., 1999. The Delphi technique as a forecasting tool: issues and analysis. *Int. J. Forecast.* 15 (4), 353–375.
- Rust, N.A., 2017. Can stakeholders agree on how to reduce human–carnivore conflict on Namibian livestock farms? A novel Q-methodology and Delphi exercise. *Oryx* 51 (2), 339–346.
- Sakamoto, F., 2023. Improving delphi study rigour with the integration of Q-methodology. *Int. J. Res. Method Educ.* 47 (2), 126–139.
- Shekelle, P.G., Kahan, J.P., Bernstein, S.J., Lucian, Leape, L.L., Kamberg, C.J., Park, R.E., 1998. The reproducibility of a method to identify the overuse and underuse of medical procedures. *N. Engl. J. Med.* 338 (26).
- Sim, J., Crookes, P., Walsh, K., Halcomb, E., 2017. Measuring the outcomes of nursing practice: a Delphi study. *J. Clin. Nurs.* 27 (1–2), e368–e378.
- Sinah, I.P., Smyth, R.L., Williamson, P.R., 2011. Using the delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. *PLoS Med.* 8 (1), e1000393. <https://doi.org/10.1371/journal.pmed.1000393>.
- Skinner, R., Nelson, R.R., Chin, W.W., Land, L., 2015. The Delphi method research strategy in studies of information systems. *Commun. Assoc. Inf. Syst.* 37 (2). <https://doi.org/10.17705/1CAIS.03702>.
- Smith, C.F., Finn, G.M., Stewart, J., McHanwell, S., 2016. Anatomical Society core regional anatomy syllabus for undergraduate medicine: the Delphi process. *J. Anat.* 228 (1), 2–14.
- Spranger, J., Homberg, A., Sonnberger, M., Niederberger, M., 2022. Reporting guidelines for Delphi techniques in health sciences: a methodological review. *Z. Evid. Fortbild. Qual. Gesundheitswes.* 172, 1–11. <https://doi.org/10.1016/j.zefq.2022.04.025>.
- Strasser, A., 2017. Delphi method variants in information systems research: taxonomy development and application. *Electron. J. Bus. Res. Methods* 15 (2), 120–133.
- Toma, C., Picioreanu, I., 2016. The delphi technique: methodological considerations and the need for reporting guidelines in medical journals. *Int. J. Public Health Res.* 4, 47–59.
- Turnbull, A.E., Dinglas, V.D., Friedman, L.A., Chessare, C.M., Sepúlveda, K.A., Bingham 3rd, C.O., Needham, D.M., 2018. A survey of Delphi panelists after core outcome set development revealed positive feedback and methods to facilitate panel member participation. *J. Clin. Epidemiol.* 102, 99–106. <https://doi.org/10.1016/j.jclinepi.2018.06.007>.
- Turoff, M., 1970. The design of a policy Delphi. *Technological Forecasting* 2 (2), 149–171.
- US Department of Health and Human Services Public Health Services Agency for Health Care Policy and Research, 1992. *Acute Pain Management: Operative or Medical Procedures and Trauma* (Rockville, MD).
- van Rijssen, L.B., Gerritsen, A., Henselmans, I., Sprangers, M.A., Jacobs, M., Bassi, C., Busch, O., Del Fernández- Castillo, C., Fong, Z.V., He, J., Jang, J.Y., Javed, A.A., Kim, S.W., Maggino, L., Mitra, A., Ostwal, V., Pellegrini, S., Shrikhande, S.V., Wilmsink, J.W., Wolfgang, C.L., van Laarhoven, H.W., Besselink, M., 2019. Core set of patient-reported outcomes in pancreatic cancer (COPRAC): an international delphi study among patients and health care providers. *Ann. Surg.* 270 (1), 158–164.
- Varndell, W., Fry, M., Lutze, M., Elliott, D., 2021. Use of the Delphi method to generate guidance in emergency nursing practice: a systematic review. *Int. Emerg. Nurs.* 56. <https://doi.org/10.1016/j.ienj.2020.100867>.
- Vogel, C., Zwolinsky, S., Griffiths, C., Hobbs, M., Henderson, E., Wilkins, E.A., 2019. Delphi study to build consensus on the definition and use of big data in obesity research. *Int. J. Obes.* 43 (12), 2573–2586.
- von der Gracht, H., 2012. Consensus measurement in delphi studies review and implications for future quality assurance. *Technol. Forecast. Soc. Change* 79, 1525–1536.
- Waggoner, J., Carline, J.D., Durning, S.J., 2016. Is there a consensus on consensus methodology? Descriptions and recommendations for future consensus research. *Acad. Med.* 91 (5), 663–668.
- Wallis, J., Burns, J., Capdevila, R., 2009. Q methodology and a delphi poll: a useful approach to researching a narrative approach to therapy. *Qual. Res. Psychol.* 6 (10).
- Weng, L.J., 2004. Impact of the number of response categories and anchor labels on coefficient alpha and test-retest reliability. *Educ. Psychol. Meas.* 64, 956.
- Yoshida, S., Rudan, I., Cousens, S., 2016. Setting health research priorities using the CHNRI method: VI. Quantitative properties of human collective opinion. *J. Glob. Health* 6, 010503.