The Delphi technique: Past, present, and future prospects — Introduction to the special issue ☆

Gene Rowe [a,*], George Wright [b]

[a] Gene Rowe Evaluations, 12, Wellington Road, Norwich NR2 3HT, UK
[b] Durham Business School, Mill Hill Lane, Durham, DH1 3LB, UK

ARTICLE INFO

ABSTRACT

The Delphi technique has been around for over half a century, so now seems a proper time to consider its past, present and future. This introduction characterises the papers in this special Delphi issue, which include both conceptual and empirical works. It summarises the main lessons that have been learned from these for the conduct of the technique, and provides a call for more and better empirical studies in the future.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

The Delphi technique has been around for some time now. It is difficult to resist the urge here to recapitulate the story of its origins, but we will: these details are noted in various papers in this special issue. One moment of history that is worth emphasising, however, is that of 1975, when the first edition of Linstone and Turoff's [1] edited book on Delphi first appeared and brought notice of the approach to a wider audience. Slowly at first, but at a seemingly growing rate, the technique has flourished, appearing in more and more academic domains and being used for more and more purposes. As evidence of the impact of that work, Google Scholar reveals that 'The Delphi Method: Techniques and Applications' has been cited over 2700 times (while it has undoubtedly been mis-cited many more times too). Furthermore, as Delphi's ubiquity has grown, so has the method evolved, with the development of numerous variants, so that it is perhaps better to talk of 'Delphi techniques' in the plural than in the singular. As such, though not quite at the 50th anniversary of publication of that seminal Delphi volume, now seems an apt moment to consider exactly how the method has developed, into what new areas and forms, and to what ends (especially since such a review/compendium has not occurred in this journal since one in 1975). Such consideration is the aim of this special issue.

2. The papers in this issue

The first paper in this issue by Roubelat and Marchais-Roubelat [2] appropriately delves the furthest back in history, to the original Oracle at Delphi. The paper does not only provide a forecasting history lesson, however: importantly, it finds parallels between the ancient approach and the modern-day namesake. Especially important is its consideration of the role of participants and the nature of expertise, as well as the issue of trust. Bolger and Wright [3] also provide a broader context for considering Delphi — although rather than historical, this focuses upon another research domain, namely social psychology. Their review and analysis discusses findings from past research that have implications for how Delphi might *work*, that is, how the method functions and how it can be enhanced. This research thus identifies a variety of important factors that need to be considered in deciding when and how to use Delphi.

☆ A special issue of technological forecasting and social change. Guest editors: Gene Rowe and George Wright.

* Corresponding author. Tel.: + 44 1603 255125.

Although this paper highlights many research questions that still need resolution, it provides a framework for conducting the necessary research — an issue to which we will return shortly.

Opposed to theoretical lessons (or lessons from other research areas), practical lessons from the conduct of 'real' Delphis are elaborated in a number of case studies. Frewer and colleagues [4] discuss the recent novel use of Delphi in the agri-food area, summarising results from a series of significant real-world exercises, in which the process was subtly amended, over the series, to take on board practical lessons learned during previous exercises. The nature of expertise and recruitment were issues of prime concern, as was the issue of language — given that many contemporary problems are global, and experts may thus speak different languages and have differing levels of access to appropriate electronic media. The issue of identifying and selecting experts is also covered in a case study by Goluchowicz and Blind [5], which looks at the area of 'standardisation'. These authors provide a thorough analysis of the recruitment problem that may act as a good model for other researchers and practitioners.

Mateos-Ronco and Server [6] also describe a novel intervention – in the area of agricultural insurance – highlighting the benefits of Delphi in terms of transparency and utility in an area that previously had limitations in these regards. And van de Linde and van der Duin [7] describe Delphi use in an extremely important new domain, namely radicalisation and terrorism. The latter case also discusses the importance of identifying both consensus *and* dissensus, emphasising an important dimension in this methodological domain. Thus, while Delphi can help demonstrate and quantify consensus, it is often where consensus is not evident that the interesting and important issues emerge. Finally here, Di Zio and Pacinelli [8] describe a truly novel approach that enables Delphi to be extended to consider spatial problems where, for example, panellists may differ in their initial views of the location of an area of greatest seismic risk.

A number of papers look at the role of Delphi, not as a standalone approach, but as a method that may be enhanced by other approaches, or that may contribute as input to others. Bañuls and Turoff [9] discuss the role of cross-impact analysis as an adjunct to Delphi in scenario development, and provide a full description to enable their approach to be replicated. On the other hand, Nowack and colleagues [10] discuss how Delphi itself might provide input into scenario planning, and they review the studies in which this has taken place. Similarly, Tapio et al [11] look at the use of disaggregative policy Delphis for scenario formation, and pose more fundamental questions about the combining of qualitative and quantitative information. They argue that 'mixed methods' are useful in order to form coherent scenarios. Landeta and colleagues [12] also describe using Delphi as part of a broader process using Nominal Group Technique and focus groups. Each of these papers importantly points to limitations to Delphi (for example, the relatively limited nature of interaction between participants) and indicates how these limitations – in terms of input and output – might be addressed through the addition of alternative techniques. Using Delphi as only part of a wider process (with qualitative and quantitative components) may well prove a means to enhance its utility; one can then avoid the question of 'which method?' and ask instead 'which methods, and in what order?' This is a more sophisticated question that may well be apt for many large, difficult, relatively well-financed real-world problems.

All of these papers – whether focused on discussing variations of Delphi, possible hybridisations, or in-depth case studies – touch upon one or more methodological issue that can impact Delphi performance/utility. Other papers in this special issue essentially start from a methodological problem per se, and from this draw implications for Delphi use more widely. Hussler et al [13] focus on the issue of panel diversity, arguing for its benefits, and discussing the issue of who is an expert and who is not. They also provide a caveat about the issue of self-interest and the potential biases this might lead to in the process, tying-in with a paper by Ecken et al [14] on the desirability of outcomes, and how a form of wishful thinking might lead to systematic bias in Delphi approaches. These latter authors suggest that a standard Delphi application may attenuate, but not totally ameliorate, the established phenomenon of desirability bias, and they provide an interesting suggestion for how Delphi might be revised to deal with bias in forecasts that are perceived as more or less desirable by individual panellists.

Bolger and colleagues [15] look at the issue of feedback and how this might impact on opinion change. In doing this, the authors follow the empirical framework already discussed by Bolger and Wright [3] and noted earlier. These authors argue that all cues to majority opinion and confidence levels of panellists should be removed from Delphi feedback in order to let strong rationales for particular panellists' opinions exert a *virtuous pull* on opinion change. Gnatzy and colleagues [16] similarly provide an empirical study that compares different Delphi formats — in this case, a conventional Delphi to a real-time version. Each of these latter two studies finds only limited differences between compared versions, which implies that the generic Delphi approach is relatively robust (i.e. that even fairly significant variations matter little) and raises questions about how to conduct and control Delphi applications in order to improve the practical efficacy of the Delphi process.

These observations lead on to two significant papers that address the crucial question of what we mean by the success of a Delphi exercise. Hasson and Keeney [17] talk about 'enhancing rigour' in Delphi, and propose reliability, validity and trustworthiness as three pertinent criteria for judging worth. They highlight the difficulties in using these criteria, particularly reliability, which has various forms, some of which may seem confounded by the precise purpose of using Delphi (e.g. to establish consensus or highlight dissensus). Parenté and Anderson-Parenté [18] also discuss the validity problem and provide one of the few retrospective attempts to validate the results from a long-term Delphi forecast (see also [19]), and indeed, possibly the longest — of events 30 years post-forecast. Although the latter authors conclude that they provide some evidence for the validity for the technique, we believe that the most important output from this paper is to demonstrate how difficult quarter-century-ahead validation is, and perhaps brings into question whether the proven accuracy of a forecast is truly a good benchmark for forecasting performance — given that desired forecasts can often be encouraged to occur, and undesired ones pre-empted, and that the interpretation of whether a particular forecast turned out to be correct can lie in large part on the interpretation of imprecise words used in the forecast formulation. This latter point brings us back to the first paper of Roubelat and Marchais-Roubelat [2], who relate the story of Croesus — a ruler who interpreted the prediction that an empire was to fall (according to the Oracle at Delphi) as being that of his enemy… when it turned out to be his own.