Available online at www.sciencedirect.com
SCIENCE DIRECT®
Technological Forecasting & Social Change 72 (2005) 377–399
Technological Forecasting and Social Change

Judgment change during Delphi-like procedures: The role of majority influence, expertise, and confidence

Gene Rowe[a,*], George Wright[b], Andy McColl[c]

[a] *Institute of Food Research, Norwich Research Park, Norwich NR4 7UA, UK*
[b] *Strathclyde Graduate Business School, University of Strathclyde, 199 Cathedral Street, Glasgow G4 OQU, UK*
[c] *Durham Business School, University of Durham, Mill Hill Lane, Durham DH1 3LB, UK*

Received 30 January 2004; received in revised form 24 March 2004; accepted 29 March 2004

Abstract

This study investigates individual opinion change and judgmental accuracy in Delphi-like groups. Results reveal that the accuracy of judgmental probability forecasts increases over Delphi rounds (in terms of proportion correct and appropriateness of confidence) when statistical summaries or written rationales are provided from other members of an individual's nominal group, but does not increase in a control iteration condition (without feedback). Additionally, subjects who gave more appropriate probability forecasts on the first round exhibited least opinion change, although measures of confidence were unrelated to opinion change. Results also show that majority opinion exerts strong opinion pull on minority opinion even when the majority favours an incorrect answer (irrespective of the nature of feedback provided). The implications of these results for the utility and conduct of the Delphi technique are discussed, in particular, with respect to selecting panellists and choosing an appropriate feedback format.

Keywords: Delphi; Forecasting; Opinion change; Majority influence

1. Introduction: the Delphi technique

The Delphi technique is a forecasting tool that was developed to allow the benefits of canvassing multiple judges without the often-corresponding deficits associated with group interaction that may arise from social processes [1–3]. It is a structured group process, in which individuals are required to give

* Corresponding author. Tel.: +44-1603-255-125.
E-mail addresses: gene.rowe@bbsrc.ac.uk (G. Rowe), wright@gsb.strath.ac.uk (G. Wright), andy.mccoll@durham.ac.uk (A. McColl).

0040-1625/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.techfore.2004.03.004

numerical judgments or forecasts over a number of rounds, with feedback being provided from the anonymous other members of the panel, and the final aggregate being taken as the process output. It is not, however, a method intended to force consensus per se—response stability rather than consensus is the signal to cease additional polling, with disagreement (as indicated by, for example, a bipolar distribution of responses) accepted as informative.

Delphi's effectiveness over comparative procedures, at least in terms of judgmental accuracy, has generally been demonstrated [3]. In a review of empirical studies of Delphi, Rowe and Wright [4] found that Delphi groups outperformed 'statistical' groups (which involve the aggregation of the judgments of noninteracting individuals) in 12 studies, underperformed these in two, and 'tied' in two others, while Delphi outperformed standard interacting groups in five studies, underperformed in one, and 'tied' in two. This trend is all the more impressive given that many laboratory studies of Delphi effectiveness have used simplified versions of the technique (e.g., with limited feedback) in simplified contexts (e.g., using nonexpert, student subjects) that might be anticipated to undermine the virtues of the technique [4,5]. We return to this issue shortly.

Although research suggests that Delphi allows improved judgment compared to alternative methods, as demonstrated in these 'technique comparison' studies, the reasons for this are still unclear, given a relative dearth of 'process' studies that have attempted to establish the precise mechanism for improvement in Delphi [5]. In this study, we attempt to advance understanding of how Delphi improves judgmental performance.

2. Research into the Delphi process: variables and measures

Generally, it is assumed that Delphi 'works' (improves judgmental accuracy) because of the feedback provided between rounds *in conjunction with its anonymity*; that is, it is assumed that the lack of cues as to the identity of the sources of the feedback allows participants to focus upon the content of that feedback rather than being distracted by extraneous social information, so being influenced by 'good' information and uninfluenced by apparently 'poor' information. Indeed, the power of feedback was demonstrated by Scheibe et al. [6], who gave 'false' feedback to panellists and found that their estimates still converged towards this fallacious figure.

However, the extent of influence of feedback is uncertain, in the sense that the variable is confounded in Delphi studies by the iteration variable. After all, it is possible that Delphi panellists may ignore feedback and still improve, by using the opportunity afforded by each new round to review their own previous estimate. Indeed, Parenté et al. [7] attempted to separate out the effects of iteration and feedback and found that, although neither iterated polling nor consensus feedback had a discernible effect upon 'if' accuracy (forecasting if an event would occur), a condition involving iteration alone resulted in improved accuracy for 'when' a newsworthy event would occur while a purely feedback condition (no iteration involved) did not. Boje and Murnighan [8] also found that accuracy decreased over rounds in a standard Delphi procedure yet improved in a purely iterative condition. These results seem to suggest that it is iteration that is the more powerful influence leading to improved accuracy, and not feedback.

Rowe and Wright [9] have argued, however, that the feedback used in previous studies has been somewhat superficial. For example, the feedback in Parenté et al. [7] simply comprised modes and medians, while that from Boje and Murnighan [8] comprised estimates and their justifications, without