



# Does the Delphi process lead to increased accuracy in group-based judgmental forecasts or does it simply induce consensus amongst judgmental forecasters? ☆

Fergus Bolger <sup>a,\*</sup>, Andrew Stranieri <sup>b</sup>, George Wright <sup>a</sup>, John Yearwood <sup>b</sup>

<sup>a</sup> Durham Business School, Durham University, UK

<sup>b</sup> Graduate School of Information Technology and Mathematical Sciences, University of Ballarat, Australia

## ARTICLE INFO

### Article history:

Received 17 January 2011

Received in revised form 3 June 2011

Accepted 6 June 2011

Available online 8 July 2011

### Keywords:

Delphi technique

Judgmental forecasting

Accuracy improvement

Consensus

Opinion change

## ABSTRACT

We investigate the relative impact of internal Delphi process factors – including panelists' degree of confidence, expertise, majority/minority positioning – and an external factor, richness of feedback – on opinion change and subsequent accuracy of judgmental forecasts. We found that panelists who had low confidence in their judgmental forecast and/or who were in a minority were more likely to change their opinion than those who were more confident and/or in a majority. The addition of rationales, or reasons, to the numeric feedback had little impact upon panelists' final forecasts, despite the quality of panelists' rationales being significantly positively correlated with accurate forecasts and thus of potential use to aid forecast improvement over Delphi rounds. Rather, the effect of rationales was similar to that of confidence: to pull panelists towards the majority opinion regardless of its correctness. We conclude that majority opinion is the strongest influence on panelists' opinion change in both the 'standard' Delphi, and Delphi-with-reasons. We make some suggestions for improved variants of the Delphi-with-reasons technique that should help reduce majority influence and thereby permit reasoned arguments to exert their proper pull on opinion change, resulting in forecast accuracy improvements over Delphi rounds.

© 2011 Elsevier Inc. All rights reserved.

## 1. Introduction

As can be seen from the papers in this Special Issue, the Delphi technique is quite frequently applied to a range of judgment problems with the expectation that judgment accuracy will be improved – relative to unstructured group judgment. However, while research supports the general conclusion that Delphi usually leads to opinion change in the direction of greater accuracy, increased accuracy is by no means guaranteed. In their review of the literature evaluating the Delphi technique, Rowe and Wright [1] found that Delphi outperformed both the statistical average of group members (by twelve studies to two with two ties) and standard interacting groups (by five studies to one with two ties) in terms of accuracy. Rowe, Wright and Bolger [2] attributed the equivocal nature of the obtained results to the highly variable formats and implementations of the Delphi technique.

The effectiveness of the Delphi technique thus appears to depend on the particular nature of its implementation – such as the number of iterations, type of feedback and size and constitution of groups. Further, the nature of the Delphi implementation interacts with the type of task that it is applied to and the characteristics of the group members (e.g. their expertise and confidence). In order to get the best improvement in judgment accuracy from applying Delphi, it would be advantageous to know the answers to these questions: what implementation of the technique works best, and under what conditions? To develop the

☆ Note that authorship order is alphabetical and therefore does not reflect the relative contribution of the authors.

\* Corresponding author at: Department of Management, Durham Business School, Mill Hill Lane, Durham DH1 3LB, UK.

E-mail address: fergus.bolger@durham.ac.uk (F. Bolger).

answers, it seems essential to have the guidance of a theoretical framework, rather than blindly trying different forms of Delphi in a variety of settings.

Rowe and Wright [3] conceptualized both judgment change and changes in judgmental accuracy as coming about through the action of both (i) internal Delphi process factors related to the individual panelist (e.g., the degree of expertise possessed by individual panelists, and (ii) each panelist's degree of confidence in a judgment) and external factors related to the particular application of the Delphi technique (e.g., the nature of feedback given to panelists between Delphi rounds and the nature of the task – either 'intellective' or 'judgmental'). Intellective tasks are those in which deduction of the already-existing truth is the focus of attention, for example. This is often the paradigm in laboratory-based evaluations of Delphi. In such settings, judgmental accuracy is easily measured. By contrast, judgmental tasks often involve forecasting – here, the judgmental forecaster can only explain and defend judgments, since the outcome has not yet occurred. Such settings are often those found in real-world Delphi applications. Wright and Rowe [4] provide a discussion of these task taxonomies.

To our knowledge, only two research studies of Delphi have been conducted within Rowe and Wright's framework: Rowe and Wright [3] and Rowe, Wright and McColl [5]. These two studies both used short-term forecasting tasks and focused on investigating the effects of increasing the richness of feedback on accuracy improvement over Delphi rounds. In many real-world applications of Delphi, the feedback procedure consists, solely, of reporting the numerical means and ranges of panelists' opinions. In the two research studies, feedback was enriched by asking judges to supply a written rationale for each judgment and, in addition, their confidence in that judgment. These studies looked also at the effects of expertise and confidence on opinion change – with the expectation that confidence and expertise would be positively related with each other and negatively related with tendency to revise opinion. Rowe et al. [5] also analyzed whether opinion change was greater when more people supported the alternative position from that chosen by the focal panelist (panelists had to predict one of two alternative outcomes).

The two studies produced conflicting findings regarding the effectiveness of providing rationales for judgments. Rowe and Wright [3] found more improvement when arguments were given, as predicted, but Rowe, Wright and McColl [5] found no advantage in eliciting and then exchanging rationales between panelists. However, a post-task examination of participants' written rationales in the latter study revealed the use of a large number of rationales that simply repeated opinions (e.g., "I think it is more than likely...") rather than causal arguments (e.g., "I think that this is more likely because..."), thus it may have been the case that most panelists were unable to provide persuasive rationales for their forecasts in this study. The two studies did, however, produce similar findings regarding the other variables: expertise and confidence were both found to be unrelated to opinion change and there was no relationship between confidence and expertise (where expertise was measured either objectively, in terms of proportion correct, or by self report).

Studies on advice taking surveyed by Bonaccio and Dalal [6] provide an indication that recommendations from an advisor to a decision maker (or judge, in these authors' terms) carry more weight if accompanied by rationales. In the, so-called, judge advisor system studies (JAS), participants are typically randomly allocated to perform the roles of either judge or advisor on a decision task. After forming a view or opinion on an issue, the judge is next presented with the views of advisors. Many JAS studies have found a bias toward favoring one's own initial opinion as judge over that of the opinions of advisors. According to Yaniv [7], this bias arises because judges have access to the full rationale underpinning their own opinion or decision and have only incomplete knowledge of advisors' rationales. In contrast, Kruger [8] contends that this bias is due to egocentricity – where a judge adheres to a default belief in the inherent superiority of his or her own judgment. In short, the extent to which discounting of advice arises because judges have access to the full rationale as opposed to egocentric bias is still an open research question.

In the current study, we compare the effects of Australia-sourced Delphi panelists exchanging rationales for their predictions relative to a standard-practice Delphi setup – which does not entail the exchange of rationales or reasons. We do this by choosing a judgmental prediction task in which many of our participants have a strong day-to-day interest – football match outcomes. According to Alomes [9], Australian Rules Football acts as a social glue in Australia – cutting across many social and gender divides and enticing many segments of the community to know a great deal about the competition. Most work places and secondary schools in the southern Australian state of Victoria have tipping competitions and a great deal of media attention is devoted to discussing forthcoming matches and analyzing the results of past matches. In this context, a number of participants are likely to be 'expert' in the judgmental prediction task that we utilized (at least relative to the rather heterogeneous set of events used by Rowe et al. [5]) and some of the panelists should, in principle, be able to provide convincing rationales for judgmental predictions of match outcomes (e.g., based on past and current form, availability of key players, home ground advantage, etc.).<sup>1</sup> In addition, we wished to investigate, in greater detail, the influence of internal Delphi process factors that might operate on opinion change. To this end, we calculated some additional measures of support for a particular predicted match outcome (tip): whether a panelist was in a minority or majority position; whether average confidence was greater for the individual panelist's tip or for the alternative tip; and, where appropriate, whether the rationale for a particular panelist's tip was stronger (and more similar, i.e. showing greater consensus with the rationales preferred by other panelists) than for the alternative tip. These new variables, plus confidence in each tip and objective expertise, were then subjected to analyses – not performed in the earlier studies – that utilized logistic regression to determine the relationship between these variables and propensity to change opinion.

<sup>1</sup> Although an empirical question, it might be expected that when all group members are highly expert there will be little or no opinion change, especially where expertise is defined as choosing the correct alternative out of the two on offer (i.e. everyone will pick the same, correct, answer). Correspondingly there should be little or no opinion change when all group members are inexpert. The greatest degree of opinion change should be observed when there is an expert minority (with well-supported and strongly-held views) and an inexpert majority (with weakly-held views).

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات