Delphi Method

Chapter · August 2016

DOI: 10.1002/9781118445112.stat07879

2 authors:

Megan M. Crawford
University of Strathclyde
6 PUBLICATIONS **3** CITATIONS
SEE PROFILE

George Wright
University of Strathclyde
167 PUBLICATIONS **6,964** CITATIONS
SEE PROFILE

Wiley StatsRef:
Statistics Reference Online

Delphi Method

By *Megan M. Grime and George Wright*

Keywords: *Delphi, group judgment, judgmental forecasting, group processes*

Abstract: This article lays out the basic structure of the Delphi method, along with common modifications, and recently developed mixed-model forecasting techniques. Delphi offers a basic structure that can be easily adapted to fit a group's needs, with respect to specific judgmental forecasting efforts.

1 Introduction

The purpose of this article is to outline and define the Delphi method and its use today. However, since its military-defense research days with the RAND Corporation, the method has been modified in a myriad of ways to better serve various uses. Therefore, a simple description is not as readily available as one would hope. However, as Linstone and Turoff note, when something can be explicitly defined, it ceases to progress[1]. As will be shown, the strength of the Delphi method is its ability to progress into new forms of application.

The simplest, most generalizable definition is offered by the RAND Corporation that originally defined Delphi as a method for "eliciting and refining group judgments" (p. v)[2]. Given a more in-depth look into the purpose of the technique and we see that the employment of the Delphi method is for facilitating structured group communication in order to gather a consensus of expert opinions in the face of complex problems, expensive endeavors, and uncertain outcomes. The principles of the method are that more minds are better than a single mind, and – when used as a forecasting tool – that structured group efforts lead to more accurate forecasts than unstructured.

Used as a forecasting technique, Delphi follows a basic structure. Anonymously, individuals offer numerical responses to a series of questions – such as the probability of an event occurring or the date-in-time that an event will have occurred. An aggregate of the responses is then generated and fed back to the group, sometimes with the reasons for the responses. Individuals are then given the option of revising (i.e., repolling) their responses on the basis of the feedback received or can restate their earlier response. The iteration and controlled feedback process continues until a predesignated stopping point is reached (i.e., number of iterations, consensus, confirmed dissensus, and stability of results). From this basic structure, modifications have been made to the method (e.g., policy Delphi[3], spatial Delphi[4], hybrid Delphi[5], and real-time Delphi[6], to name a few).

Note that even without any repolling, simply utilizing the median of a group's opinions on, say, the unit sales of a new product in the first year of production will provide more accuracy than that due to at least 50% of the individual panelists. *With* repolling and feedback, it is assumed that the median response of the group shifts nearer to the true value of the outcome to be predicted. Improvement is thought to result

University of Strathclyde Business School, Glasgow, UK

DOI: 10.1002/9781118445112.stat07879

from (i) opinion change in "swingers" – who change their less-firmly grounded opinions after receiving feedback about the opinions of other, anonymous, panelists and (ii) opinion stability in "holdouts" – who are assumed to be more accurate than "swingers"[2]. Results generally suggest that Delphi groups are more accurate than traditional groups. In a review of research, Rowe and Wright found that Delphi groups outperformed traditional groups by a score of five studies to one, with two ties, and with one study showing task-specific support for both techniques[7].

2 When to Use Delphi

Several studies have revealed a selection of criteria that help determine when it is best to use Delphi[1,8–10]. The Delphi method is beneficial when the problem at hand can benefit from collective, subjective judgments or decisions and when group dynamics do not allow for effective communication (e.g., time differences, distance, and personality conflicts). Physical group meetings may be too costly. However, even when face-to-face group meetings are possible, a forecasting effort may benefit from the supplement of a Delphi process. This is also true when using other (futures) methodologies that can benefit from the addition of a Delphi process to aggregate varied individual opinions (e.g., scenario planning, cross-impact analysis and interpretive structural modeling[11], and Thurstone's Law scaling method[12]).

3 How to Use Delphi

3.1 Investigator

The investigator (i.e., researcher, facilitator, or moderator) is a key element to the Delphi method. It is the responsibility of the investigator to understand the project at hand, employ the appropriate tools (whether Delphi alone or in conjunction with other techniques), recruit experts, know the effectiveness of the group size and composition, gather appropriate knowledge and opinions from the experts, facilitate their creative thinking as individuals and as a group, collect, compose, synthesize, and redistribute their varied responses (both qualitative and quantitative), and do this all in an effective, timely, and objective manner[1]. Therefore, the skill set the investigator brings to the Delphi method is as important as all other aspects of the method itself. Unfortunately, fewer studies focus on the requirements of the investigator as opposed to the technique itself. This is, perhaps, something that should not be so overlooked in the literature, given what we now know about cognitive heuristics and biases commonly employed in the decision-making process.

One investigator bias already widely recognized as a potential problem within the Delphi method comes from inserting his/her own preconceived opinions into the synthesis and feedback process. With respect to quantitative evaluations, Black *et al*.[13] suggest one way to help counter, but not fully eliminate, investigator bias is to use standardized checklists or rubrics for grading of responses. The rounds of a Delphi process response may also involve the feedback of qualitative data (i.e., rationales for individual assessments). There are no firm guidelines on which qualitative summation method to use. Green *et al*.[14] support less active intervention in reducing and categorizing statements. This increases transparency but leaves lengthy, detailed feedback. Mixing both qualitative and standardized quantitative measures may reduce potential bias in the synthesis requirement of the investigator's efforts.

3.2 Experts

There are several suggestions on how best to select a group of experts. Although each situation and pool of experts from which to select will be unique, some general principles for the process are offered.

Wiley StatsRef: Statistics Reference Online, © 2014–2016 John Wiley & Sons, Ltd.

DOI: 10.1002/9781118445112.stat07879

Rowe and Wright[7] summarize the following principles for using expert opinion in applications of Delphi:

- Use experts with appropriate domain knowledge[15,16].
- Use heterogeneous experts[10].
- Use between 5 and 20 experts[17].
- For Delphi feedback, provide the mean or median estimate of the panel plus the rationales from all panelists for their estimates[18,19].
- Continue Delphi polling until the responses show stability. Generally, three structured rounds are enough[20,21].
- Obtain the final forecast by weighting all the experts' estimates equally and aggregating them.

3.3 Design

The method should be designed in such a way that the questions elicit the experts' reasoning behind their individual responses. Not only does the act of offering supportive reasoning help confirm the plausibility and strength of the expert's response to him/herself, it also allows for better evaluation efforts by the other experts when presented with the aggregate of responses in the next feedback and iteration. Further information can be gathered at this stage of the process, such as perceived relevant factors concerning the focal Delphi assessment, best estimates of these factors, as well as any information missing that the expert feels would improve his/her confidence in responding to the focal Delphi assessment[9].

3.4 Assessments and Controlled Feedback

First-round estimates are generally the most divergent of all the questionnaire rounds[2]. It is the investigator's key responsibility to collect, edit, and synthesize responses and then return a statement that reflects an aggregate of the group's position. After receiving the feedback, the experts are given a chance to reevaluate their original responses along with the new aggregate information with the opportunity to add, alter, or maintain their individual responses. Each group will determine how many iterations they go through before reaching a stopping point. This is usually determined ahead of time (i.e., number of iterations, consensus reached, and stability of results). The goal is to come to either a strong consensus or a continuing lack of consensus. The latter indicting strong, self-justified differences in opinion.

It is important to note that many studies have shown consensus can often be influenced by the aggregate report after each iteration. That is, the mean, median, or majority opinion – given as feedback between Delphi rounds – can have more of an effect on opinion formation/stability than accuracy[22,23]. Suggestions of how to avoid such biased consensus include *facilitated information exchange* that enables experts to review other group members' qualitative reasoning in an effort to enhance recognition of justification errors. As such, markings of majority viewpoints or measures of panellist confidence should be removed from quantitative feedback between Delphi rounds. A recent meta-analysis of six Delphi studies revealed that desirability of a forecast can also affect probability estimation over Delphi rounds[24]. These investigators advocated the use of a posthoc procedure to better identify those judgmental forecasts that are more likely to be affected by a desirability bias.

The anonymity in the iteration process allows participants to avoid common socially induced bias, like appearing consistent in one's opinions and appearing committed to socially desirable future events. These are social factors that operate largely at the implicit level. When an expert's declaration is made public, the desire to appear consistent in thinking and then to commit to that opinion are strengthened[25]. As noted in an early paper by Dalkey, Brown, and Cochran, the Delphi method allows (and indeed encourages)

DOI: 10.1002/9781118445112.stat07879

one's prior opinions from previous rounds to be reexamined in later rounds, in conjunction with the new aggregate of data and reasoning[2].

4 Mixed Methods Approaches

Integrating different futures tools may be beneficial. Landeta, Barrutia, and Lertxundi created a Hybrid Delphi method by integrating elements of focus groups, nominal groups, and Delphi techniques in a chronological and complimentary manner[5]. Through a process of synthesis, the technique uses sequential ordering to tailor each step to the overall objective and context. The goal of the design is to use this integrative technique in a manner where all three methods each resolve the recognized problems inherent in the other methods. Varho and Tapio offered a new methodology using the systematic efforts of Delphi to integrate expert opinions into a tabular format. This information – both qualitative and quantitative – then provided a more methodical presentation for the broad, intuitive scenario process[26]. Warth *et al*. used Delphi with different representative stakeholder groups to inform scenario development on the future of electric vehicles. Projections that reached consensus were used for scenario development, those that did not were used to create a variety of scenario storylines[27]. Pincombe *et al*. suggested a model with an on-line Delphi process to form the dimensions of later scenario processes. Through an iteration process, consensus is reached on key driving forces, with the consequence of lessening the time commitments of the experts[28]. Heiko and Stillings used a cross-national Delphi process to garner further evaluations on projections in the early stages of innovation management. Again, consensus informed later scenarios[29].

5 Comparison of Methods and Tools

The nominal group technique has experts record their suggestions and ideas, which are then compiled by a investigator for all to see. Similar to the Delphi method, discussions and idea-generating can go through several rounds before a stopping point is reached. However, unlike Delphi, the experts then anonymously rank the final selection of ideas and the investigator summarizes the ranks[30,31].

Decision conferencing follows a process similar to Delphi. Experts share insights, judgments, and suggestions on a particular issue. A decision-analysis model is then synthesized by a decision analyst to produce, for example, a decision tree analysis from all the judgments. This model is then presented to the group, along with the consequences. An iteration process of model/consequences and reassessment continues until the results of the analysis are accepted by the group membership. The final stage produces a report that summarizes the results, and when pertinent, includes decision recommendations. Where this process differs from Delphi is in the lack of anonymity. The decision-conference process assumes that work is more effective if all the participants are in the same physical space (e.g., meeting room) to directly communicate with one another[32].

When the investigator must synthesize and average responses from a group of experts – as in producing the Delphi yield – responses are usually given equal weight. This is, however, not the only way to combine individual judgments. The linear opinion pool technique uses a differential weighting system[33]. These weights can be assigned by the investigator or by the experts themselves. As well, the weighting system can be determined in a number of ways, including a measure of individual experts' prior performance[34], pregroup testing or questionnaires[35], or even as a relation to the distance from one's own focal judgmental assessment[36].

Group-based judgments that are not the result of a structured group process may be produced by processes that suppress minority-based opinion (i.e., groupthink)[37]. This is why it is important for both the experts and the investigator to understand the dynamics of unstructured group-based process in order

Wiley StatsRef: Statistics Reference Online, © 2014–2016 John Wiley & Sons, Ltd.

DOI: 10.1002/9781118445112.stat07879

to employ the right kind of judgment–elicitation tools. Empirical work in the area of social dilemma games show how individuals' natural decisions may possess, or lack, cooperation – leading to both optimal and suboptimal choices[38,39]. Descriptive tools, like a decision scheme matrix, can be used to help the group members decide how they want to proceed through a group-based exercise toward an eventual consensus[40]. For example, individual group members may choose to maintain, or not maintain, a judgment if there is at least one supporting voice, versus a majority voice, versus a unanimous decision.

6 Closing Discussion

Arguably, the biggest element impacting the effectiveness of the Delphi process is the human element. The process relies on human judgment throughout the entirety of the process. As has been shown in extensive psychology research, human judgment – both at the individual and group-based levels – may not be free from the impact of cognitive biases[41]. Recall that the Delphi process is, in essence, a methodology for the structured aggregation of individual judgments. This does not mean any and all applications of the Delphi method may produce poor quality aggregated judgments and forecasts. However, it does mean that both practitioners and participants should be aware of such influencing phenomenon.

Related Articles

Decision Conferencing/Facilitated Workshops; **Expert Judgment**; **Group Decision**; **Game Theory**; **Nash Equilibrium**.

03600
03603
03853
00060
00103

References

[1] Linstone, H.A. and Turoff, M. (2002) *The Delphi Method: Techniques and Applications*, *http://www.is.njit.edu/pubs/delphibook/index.html* (accessed 02 February 2016).

[2] Dalkey, N.C., Brown, B.B., and Cochran, S. (1969) *The Delphi Method: An Experimental Study of Group Opinion*, RAND Corporation, Santa Monica, CA.

[3] Turoff, M. (1970) The design of a policy Delphi. *Technol. Forecast. Soc. Change*, **2**, 149–171.

[4] Zio, S.D. and Pacinelli, A. (2011) Opinion convergence in location: a spatial version of the Delphi method. *Technol. Forecast. Soc. Change*, **78**, 1565–1578.

[5] Landeta, J., Barrutia, J., and Lertxundi, A. (2011) Hybrid Delphi: a methodology to facilitate contribution from experts in professional contexts. *Technol. Forecast. Soc. Change*, **78**, 1629–1641.

[6] Hartman, F.T. and Baldwin, A. (1995) Using technology to improve Delphi method. *J. Comput. Civil Eng.*, **9**, 244–249.

[7] Rowe, G. and Wright, G. (2001) Expert opinions in forecasting: role of the Delphi technique, in *Principles of Forecasting: A Handbook for Researchers and Practitioners* (ed J.S. Armstrong), Kluwer Academic Publishers, Norwell, MA.

[8] Bolger, F. and Wright, G. (2011) Improving the Delphi process: lessons from social psychological research. *Technol. Forecast. Soc.Change*, **78**, 1500–1513.

[9] Dalkey, N. and Helmer, O. (1963) The experimental application of the Delphi method to the use of experts. *Manag. Sci.*, **9**, 458–467.

[10] Delbecq, A.L., Van de Ven, A.H., and Gustafson, D.H. (1975) *Group Techniques for Program Planning: A Guide to Nominal and Delphi Processes*, Scott Foresman and Co, Glenview, IL.

[11] Bañuls, V.A. and Turoff, M. (2011) Scenario construction via Delphi and cross-impact analysis. *Technol. Forecast. Soc. Change*, **78**, 1579–1602.

[12] White, C., Turoff, M., Van de Walle, B. (2007) *A Dynamic Delphi Process Utilizing a Modified Thurstone Scaling Method*. Collaborative Judgment in Emergency Response Proceedings of the 4th Annual Information Systems on Crisis and Response Management, ISCRAM, Delft, Netherlands, vol. 4, pp. 13–16.

[13] Black, N., Murphy, M., Lamping, D., *et al.* (1999) Consensus development methods: a review of best practice in creating clinical guidelines. *J. Health Serv. Res. Policy*, **4**, 236–248.

DOI: 10.1002/9781118445112.stat07879

[14] Green, B., Jones, M., Hughes, D., and Williams, A. (1999) Applying the Delphi technique in a study of GPs information requirements. *Health Soc. Care Comm.*, **7**, 198–205.
[15] Hora, S.C. (2014) Expert judgment, in *Wiley StatsRef: Statistics Reference Online* (eds P. Brandimarte, B. Everitt, G. Everitt, W. Piegorsch, and F. Ruggeri), John Wiley & Sons, Ltd, Chichester.
[16] Jolson, M.A. and Rossow, G. (1971) The Delphi process in marketing decision making. *J. Market. Res.*, **8**, 443–448.
[17] Armstrong, J.S. (1985) *Long Range Forecasting: From Crystal Ball to Computer*, 2nd edn, John Wiley & Sons, Inc., New York.
[18] Rowe, G., Wright, G., and Bolger, F. (1991) The Delphi technique: a reevaluation of research and theory. *Technol. Forecast. Soc. Change*, **39**, 235–251.
[19] Rowe, G. and Wright, G. (1996) The impact of task characteristics on the performance of structured group forecasting techniques. *Int. J. Forecast.*, **12**, 73–89.
[20] Brockhoff, K. (1975) The performance of forecasting groups in computer dialogue and face to face discussions, in *The Delphi Method: Techniques and Applications* (eds H. Linstone and M. Turoff), Addison-Wesley, London.
[21] Rohrbaugh, J. (1979) Improving the quality of group judgment: social judgment analysis and the Delphi technique. *Organ. Behav. Hum. Perform.*, **24**, 73–92.
[22] Bolger, F., Stranieri, A., Wright, G., and 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. Change*, **78**, 1671–1680.
[23] Kerr, N.L. and Tindale, R.S. (2001) Group-based forecasting? A social psychological analysis. *Int. J. Forecast.*, **27**, 14–40.
[24] Ecken, P., Gnatzy, T., and von der Gracht, H.A. (2011) Desirability bias in foresight: consequences for decision quality based on Delphi results. *Technol. Forecast. Soc. Change*, **78**, 1654–1670.
[25] Cialdini, R. (2007) *Influence: The Psychology of Persuasion*, William Morrow and Co, New York.
[26] Varho, V. and Tapio, P. (2013) Combining the qualitative and quantitative with the Q2 scenario technique – the case of transport and climate. *Technol. Forecast. Soc. Change*, **80**, 611–630.
[27] Warth, J., Heiko, A., and Darkow, I.L. (2013) A dissent-based approach for multi-stakeholder scenario development – the future of electric drive vehicles. *Technol. Forecast. Soc. Change*, **80**, 566–583.
[28] Pincombe, B., Blunden, S., Pincombe, A., and Dexter, P. (2013) Ascertaining a hierarchy of dimensions from time-poor experts: linking tactical vignettes to strategic scenarios. *Technol. Forecast. Soc. Change*, **80**, 584–598.
[29] Heiko, A. and Stillings, C. (2013) An innovation-focused scenario process – a case from the materials producing industry. *Technol. Forecast. Soc. Change*, **80**, 599–610.
[30] Delbecq, A.L., Van de Ven, A.H., and Gustafson, D.H. (1975) *Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes*, Foresman Scott, Glenview, IL.
[31] Van de Ven, A.H. and Delbecq, A.L. (1974) The effectiveness of nominal Delphi, and interacting group decision making processes. *Acad. Manage. J.*, **17** (4), 605–621.
[32] Phillips, L.D. (2014) Decision conferencing/facilitated workshops, in *Wiley StatsRef: Statistics Reference Online* (eds P. Brandimarte, B. Everitt, G. Everitt, W. Piegorsch, and F. Ruggeri), John Wiley & Sons, Ltd, Chichester.
[33] Genest, C. and McConwy, K. (1990) Allocating the weights in the linear opinion pool. *J. Forecast.*, **9**, 53–73.
[34] Stone, M. (1961) The linear opinion pool. *Ann. Math. Stat.*, **32**, 1339–1342.
[35] Cooke, R.M. (1991) *Experts in Uncertainty*, Oxford University Press, Oxford.
[36] Carvalho, A. and Larson, K. (2013) *A Consensual Linear Opinion Pool*. Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence. AAAI Press.
[37] Morton, A. (2014) Group decision, in *Wiley StatsRef: Statistics Reference Online* (eds P. Brandimarte, B. Everitt, G. Everitt, W. Piegorsch, and F. Ruggeri), John Wiley & Sons, Ltd, Chichester.
[38] Lucas, W.F. (2014) Game theory, in *Wiley StatsRef: Statistics Reference Online*, John Wiley & Sons, Ltd, Chichester.
[39] Coughlin, P.J. (2014) Nash equilibrium, in *Wiley StatsRef: Statistics Reference Online* (eds P. Brandimarte, B. Everitt, G. Everitt, W. Piegorsch, and F. Ruggeri), John Wiley & Sons, Ltd, Chichester.
[40] Davis, J.H. (1973) Group decision and social interaction: a theory of decision schemes. *Psychol. Rev.*, **80**, 97–125.
[41] Kahneman, D., Slovic, P., and Tversky, A. (1982) *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge.

Wiley StatsRef: Statistics Reference Online, © 2014–2016 John Wiley & Sons, Ltd.

DOI: 10.1002/9781118445112.stat07879