

Effective Use of the Delphi Process in Research:
Its Characteristics, Strengths and Limitations ¹

Jerry A. Somerville

2008

The Delphi method was first developed in the 1950s by Olaf Helmer, Nicholas Rescher, Norman Dalkey, and others at the RAND corporation (Gordon, 1994). The intent of the Delphi, as it was originally conceived, was to create a method, using expert opinions, to forecast long-range trends related to the military potential of future science and technology and their effects on political issues (Gordon, 1994; Linstone & Turoff, 1975).

Delphi Characteristics

A Delphi is an iterative process, normally three to four rounds, involving a series of questionnaires, each building on the results of the previous one. The results of each round are compiled and returned to the participants. Over successive iterations, participants are able to reevaluate their responses in light of the compiled responses of all participants. Responses to the questionnaires are made anonymously. Participants are known to the researchers but not necessarily to the other participants. The anonymity of panelists enhances the probability that opinions are considered in and of themselves without being influenced by the person who expressed the opinions.

The Delphi is most appropriately used when the “primary source of information sought is informed judgment” (Ziglio, 1996, p. 21). “The value of the Delphi method rests with the ideas it generates, both those that evoke consensus and those that do not. The arguments for the extreme positions also represent a useful product” (Gordon, 1994, p. 4). The concepts of

Excerpt from Somerville, J. A. (2007). *Critical factors affecting the meaningful assessment of student learning outcomes: A Delphi study of the opinions of community college personnel*. Unpublished doctoral dissertation, Oregon State University, Corvallis, OR.

developing and understanding a subject and the fact that participants possess that knowledge are central to the Delphi method (Gordon, 1994; Linstone & Turoff, 1975) and represent key features in qualitative research (Creswell, 2002).

Sackman (1974) identified the following as the characteristics of a conventional Delphi:

- A formal and structured questionnaire is used.
- Questionnaire items may be generated by the moderator, the panelists, or both.
- Either quantitative or qualitative scales may be used.
- The process consists of two or more rounds.
- Questionnaires may or may not include open-ended questions.
- Feedback from each round is in the form of statistical feedback, usually involving some measure of central tendency and some measure of dispersion.
- Feedback from each round may include selected textual information.
- Individual responses to items are kept anonymous.
- Iteration with feedback continues until consensus is reached, as determined by the moderator.
- Participants do not meet face to face and may be geographically dispersed.
- Outliers (i.e. upper and lower quartile) may be asked to justify their responses in writing.

The Delphi is a flexible method built on four basic features: “structured questioning, iteration, controlled feedback, and anonymity of responses” (Lang, 1995, p. 3). Advantages of the method are that information can be gathered from a geographically diverse panel of participants; that panelists have anonymity, which reduces the halo effects associated with the

Excerpt from Somerville, J. A. (2007). *Critical factors affecting the meaningful assessment of student learning outcomes: A Delphi study of the opinions of community college personnel*. Unpublished doctoral dissertation, Oregon State University, Corvallis, OR.

opinions of prominent participants; and that panelists have time to consider carefully their responses before replying (Adams & O'Brien, 2004; Garrod, 2004; Gordon, 1994).

Some disadvantages until recently were the time and expense of designing paper and pencil questionnaires, mailing surveys, compiling responses, and following up with non-respondents for multiple iterations of the process. These disadvantages are resolved with the use of an electronic version of the Delphi method, called the "e-Delphi." The time and expense of the process are dramatically reduced, data are electronically compiled, and more detailed information can be returned to participants (Chou, 2002; Human-Environment Regional Observatory, 2001).

Another disadvantage of Delphi is a potentially high attrition rate. Because the method requires lengthy responses in the early rounds of the process and the active participation of panelists over several weeks, the potential for a high drop-out rate of panelists exists (Borg & Gall, 1983). Several steps can be taken to mitigate attrition: minimizing frustration through ease of access to and navigation of the survey Website, communicating clearly to panelists about the extent of their expected involvement, providing speedy feedback of the results of each round, and encouraging non-responders to respond through systematic follow-up contacts.

Sackman (1974), the most prominent critic of the method, concluded that Delphi studies were poorly administrated, were unscientific, and did not conform to standard psychometric principles. However, others have challenged Sackman's criticism. Linstone (1975) believed that Sackman simply missed the point of the Delphi, stating it was not a method for determining causality but a technique to facilitate deliberation on a problem and to aggregate the informed opinions of experts. The Delphi was proposed as an alternate paradigm to the "tradition-bound...objective" (Linstone, 1975, p. 559) attitude of the time and should not be judged by

Excerpt from Somerville, J. A. (2007). *Critical factors affecting the meaningful assessment of student learning outcomes: A Delphi study of the opinions of community college personnel*. Unpublished doctoral dissertation, Oregon State University, Corvallis, OR.

conventional experimental and psychometric standards. Ziglio (1996) stated that there is no reason why the Delphi method should be considered any “less methodologically robust than techniques such as interviewing, case analysis, or behavioral simulations” (p. 13). The Delphi method has also been criticized for lacking standards for determining who was an expert (Fischer, 1978; Mitchell, 1991; Rowe & Wright, 1999; Sackman, 1974; Stewart, 1987), lacking a common starting point that provided panelists with current assumptions and findings (Fischer, 1978), facilitating conformity rather than consensus (Sackman, 1974; Stewart, 1987), promoting quick answers to complex problems, and suppressing divergent views (Stewart, 1987). According to Rowe and Wright (1999), panelists with divergent views were more likely to drop out, and thus consensus may be the result of attrition.

Feedback to panelists in the form of reasons for their ratings has been shown to improve the accuracy of group judgments. However, according to Rowe and Wright (1999), feedback of reasons or rationales behind panelists’ estimates has been rare in Delphi studies. They added, “Since it is through the medium of feedback that information is conveyed from one panelist to the next, by limiting feedback one also limits the scope of panelists’ aggregate accuracy” (p. 369).

Delphi Panels

The size of Delphi panels can vary widely and there is disagreement about what constitutes an appropriate panel size. Clayton (1997) indicated that by rule of thumb 15 to 30 people is the norm for homogeneous groups (e. g., professors from the same discipline), where Ziglio (1996) reported that 10 to 15 people produce good results in a homogeneous panel. For heterogeneous groups (people with expertise on a topic but from different social or professional groups), Clayton (1997) reported that only 5 to 10 experts are needed. Gordon (1994) indicated

Excerpt from Somerville, J. A. (2007). *Critical factors affecting the meaningful assessment of student learning outcomes: A Delphi study of the opinions of community college personnel*. Unpublished doctoral dissertation, Oregon State University, Corvallis, OR.

that most Delphi studies use panels of 15 to 35 people. However, in two separate studies investigating the size of Delphi panels, no consistent relationship between panel size and effectiveness criteria was found (Rowe & Wright, 1999). In other literature on aggregating group opinions, groups of 6 to 12 members were determined to be optimum (Hogarth, 1978; Mitchell, 1991). Also, it was found that the more the members differed, the larger the group should be. Further, all things being equal, the larger the group, the more reliable their aggregate judgment will tend to be. However, beyond group sizes of 20 to 25, there were only minimal improvements in reliability (Hogarth, 1978).

It appears that panels of experts who also have a diversity of perspectives produce more accurate judgments than experts who are more homogeneous (Lang, 1995; Powell, 2003; Wallsten, Budescu, Erev, & Diederich, 1997). Winkler and Poses (1993) demonstrated that a group of physicians representing different specialties were better at predicting the survival of patients admitted to intensive care units than were individual physicians or a group of physicians representing the same specialty. Thus, the accuracy of the aggregated group opinions could be improved by selecting panelists, for example, who are knowledgeable about student learning outcomes assessment and who represent various campus constituencies (i.e., administrator, campus researcher, instructional faculty member, student services personnel).

Criteria for Truth

Linstone and Turoff (1975) suggested that the Delphi method is appropriate for problems that do not lend themselves “to precise analytical techniques but [could] benefit from subjective judgments on a collective basis” (p. 4). Delphi is a method for structuring a group communication process to systematically explore and gain insight into a problem (Mitroff & Turoff, 1975; Sackman, 1974). It is founded on the belief that collecting data precedes the

Excerpt from Somerville, J. A. (2007). *Critical factors affecting the meaningful assessment of student learning outcomes: A Delphi study of the opinions of community college personnel*. Unpublished doctoral dissertation, Oregon State University, Corvallis, OR.

development of theory (Mitroff & Turoff, 1975). In a report on a study to develop a framework for evaluating qualitative research, Spencer, Richie, Lewis, and Dillon (2003) stated that one form of truth in research is “agreement that it is true (a consensus view of truth)” (p. 62). This is the case for a Delphi. Truth is experiential, derived inductively, and based on “sufficient widespread agreement ... by a group of ‘experts’” (Mitroff & Turoff, 1975, p. 21). Scheele (1975) explained that in the Delphi process, reality is negotiated by the group. It is constructed through the perceptions the participants bring to the discussion.

Because the number of respondents is usually small, Delphis do not, and are not intended to, produce statistically significant results; in other words, the results provided by any panel do not predict the response of a larger population or even a different Delphi panel. They represent the synthesis of the opinions of the particular group, no more, no less. (Gordon, 1994, pp. 3-4)

Aggregating Panelists’ Opinions

The Delphi method is based on panelists achieving consensus; however, there is no standard method for determining consensus (Hasson, Keeney, & McKenna, 2000; Mitchell, 1991). According to Mitchell (1991), “a growing body of research questions consensus as a stopping criterion” (p. 347). It appears that the most change in panelists’ responses occurs within the first two rounds and that not much is gained in further iterations (Mitchell, 1991).

One method for aggregating the subjective judgments of panelists to produce a collective opinion is by simply averaging participant responses. This has been shown to be a robust method for aggregating a group’s judgment (Clemen, 1989; Clemen & Winkler, 1986; Larrick & Soll, 2003; Wallsten et al., 1997; Winkler & Clemen, 2004). Further, when using a rating scale “the reliability of ratings can be greatly improved by pooling the results from several judges who have made their ratings independently” (Helmstadter, 1964, p. 198).

Excerpt from Somerville, J. A. (2007). *Critical factors affecting the meaningful assessment of student learning outcomes: A Delphi study of the opinions of community college personnel*. Unpublished doctoral dissertation, Oregon State University, Corvallis, OR.

It appears that group judgment is improved when the members received text as well as statistical feedback. Rowe and Wright (1999) reported on studies that compared “reasons” or “rationale” feedback to statistical feedback (i.e., mean, median). They concluded that there was greater improvement in the accuracy of judgments in successive rounds of a Delphi when panelists’ reasons or rationales for their ratings were given to other panelists in the group as opposed to panelists receiving only statistical feedback.

Ensuring Soundness of Data

Delphi is one form of survey procedure and, as such, should follow the guidelines for good survey design and administration. Ensuring accurate results depends on quality control throughout the administration of the process (Scheuren, 2004). This includes maximizing respondent motivation to participate, ensuring the clarity of the questions and respondent instructions (Barribeau et al., 2005), devising a plan to follow up on non-respondents, pre-testing the questionnaire and survey procedures, coding the survey information accurately, and recording the data correctly (Scheuren, 2004).

Since the Delphi method was first introduced in the 1950s, much has been learned that can improve its rigor and in turn ensure the quality of the data. The suggestions center around three themes: panel selection and motivation, questionnaire construction, and process management:

Panel selection and motivation suggestions indicate:

- Select panel members based on knowledge of the issue and diversity of perspective (Garrod, 2004; Lang, 1995; Linstone & Turoff, 1975; Ziglio, 1996).
- Provide enough incentive to maintain panelists’ motivation to persist to the conclusion of the study (Garrod, 2004; Uhl, 1983).

Excerpt from Somerville, J. A. (2007). *Critical factors affecting the meaningful assessment of student learning outcomes: A Delphi study of the opinions of community college personnel*. Unpublished doctoral dissertation, Oregon State University, Corvallis, OR.

- Ensure that panelists feel that their contributions are valued (Mitchell, 1991; Turoff & Hiltz, 1996).
- Communicate to panelists that they are members of a group with similar expertise to theirs (Krebsbach, 1998).

Questionnaire construction suggestions are:

- Provide clear written instructions to panel members (Hasson et al., 2000; Ziglio, 1996).
- Make questionnaire statements clear, concise, free of ambiguities, and easily understood by panelists from varied backgrounds (Garrod, 2004; Lang, 1995; Uhl, 1983).
- Pre-test the questionnaire (Lang, 1995; Mitchell, 1991).
- Take care to keep the intent of panelist responses intact when reporting responses back to other panel members (Lang, 1995).

Process management suggestions include:

- Guard against imposing the biases of the moderator on panel members (Uhl, 1983).
- Provide panelists with a brief account of the origin and purpose of the study (Lang, 1995).
- Establish the credibility of the research and the researcher (Mitchell, 1991).
- “Allow enough time between rounds to prepare and distribute feedback, but do not allow so much time that panelist lose interest” (Uhl, 1983, p. 88).
- Acknowledge divergent opinions (Gordon, 1994; Lang, 1995; Uhl, 1983).
- Consider Delphi results in light of the results from other methods (Lang, 1995).

Excerpt from Somerville, J. A. (2007). *Critical factors affecting the meaningful assessment of student learning outcomes: A Delphi study of the opinions of community college personnel*. Unpublished doctoral dissertation, Oregon State University, Corvallis, OR.

Summary

The Delphi is appropriate to use when the researcher seeks the informed judgment of participants who have knowledge of a particular topic. One way to reveal that knowledge is through a structured group participation activity. With the Delphi method, participants are able to present and rationalize their opinions about the topic being researched. Also they have the opportunity to consider the opinions of others, reconsider their own opinions, and assess the relative importance of each opinion presented. To insure soundness of data researchers should pay particular attention to panel selection and motivation, questionnaire construction, process management, and the method for aggregating panelists' opinions.

Excerpt from Somerville, J. A. (2007). *Critical factors affecting the meaningful assessment of student learning outcomes: A Delphi study of the opinions of community college personnel*. Unpublished doctoral dissertation, Oregon State University, Corvallis, OR.

References

- Adams, C., & O'Brien, H. (2004). *Delphi method*. Retrieved December 12, 2004, 2004, from www.arches.uga.edu/~cadamrun/7320/delphi.htm
- Barribeau, P., Butler, B., Corney, J., Doney, M., Gault, J., Gordon, J., et al. (2005). *Survey research*. Retrieved May 23, 2005, from <http://writing.colostate.edu/references/research/survey>
- Borg, W. R., & Gall, M. D. (1983). *Educational research: An introduction* (4th ed.). New York: Longman.
- Chou, C. (2002). Developing the e-Delphi system: A web-based forecasting tool for educational research. *British Journal of Educational Technology*, 33(2), 233-236.
- Clayton, M. J. (1997). Delphi: A technique to harness expert opinion for critical decision-making tasks in education. *Educational Psychology*, 17(4), 373-384.
- Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5, 559-583.
- Clemen, R. T., & Winkler, R. L. (1986). Combining economic forecasts. *Journal of Business and Economic Statistics*, 4(1), 39-46.
- Creswell, J. W. (2002). *Educational research: Planning conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson Education.
- Fischer, R. G. (1978). The Delphi method: A description, review, and criticism. *Journal of Academic Librarianship*, 4(2), 64-70.
- Garrod, B. (2004). *Lecture notes on the Delphi technique*. Retrieved December 12, 2004, from <http://users.aber.ac.uk/bgg/rm.htm>
- Gordon, T. J. (1994). *The Delphi method*. Washington, DC: American Council for the United Nations University.
- Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, 32(4), 1008-1015.
- Helmstadter, G. C. (1964). *Principles of psychological measurement*. New York: Appleton-Century-Crofts.
- Hogarth, R. M. (1978). A note on aggregating opinions. *Organizational Behavior and Human Performance*, 21, 40-46.
- Human-Environment Regional Observatory. (2001). *The HERO e-Delphi system: Overview and implementation*. Retrieved December 22, 2004, from http://hero.geog.psu.edu/products/Delphi_white_paper.pdf
- Krebsbach, S. G. (1998). *A set of learning outcomes for the learner in the two year institution of higher education, in order to function in work, community, and family at the beginning of the 21st century*. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, MN.
- Lang, T. (1995). *An overview of four futures methodologies*, from <http://www.futures.hawaii.edu/jrnls.html>
- Larrick, R. P., & Soll, J. B. (2003). *Intuitions about combining opinions: Misappreciation of the averaging principle*. Retrieved March 25, 2006, from <http://knowledge.insead.edu/doc.cfm?cd=1265>
- Linstone, H. A. (1975). Eight basic pitfalls: A checklist. In H. A. Linstone & M. Turoff (Eds.), *The Delphi method: Techniques and applications*. Boston, MA: Addison-Wesley.
- Excerpt from Somerville, J. A. (2007). *Critical factors affecting the meaningful assessment of student learning outcomes: A Delphi study of the opinions of community college personnel*. Unpublished doctoral dissertation, Oregon State University, Corvallis, OR.

- Linstone, H. A., & Turoff, M. (Eds.). (1975). *The Delphi method: Techniques and applications*. Boston, MA: Addison-Wesley.
- Mitchell, V. W. (1991). The Delphi technique: An exposition and application. *Technology Analysis & Strategic Management*, 3(4), 333-358.
- Mitroff, I. I., & Turoff, M. (1975). Philosophical and methodological foundations of Delphi. In H. A. Linstone & M. Turoff (Eds.), *The Delphi method: Techniques and applications* (pp. 17-34). Boston, MA: Addison-Wesley.
- Powell, C. (2003). The Delphi technique: Myths and realities. *Journal of Advanced Nursing*, 41(4), 376-382.
- Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: Issues and analysis. *International Journal of Forecasting*, 15, 353-375.
- Sackman, H. (1974). *Delphi assessment: Expert opinion, forecasting and group process*. Santa Monica, CA: Rand Corporation.
- Scheele, D. S. (1975). Reality construction as a product of Delphi interaction. In H. A. Linstone & M. Turoff (Eds.), *The Delphi method: Techniques and applications* (pp. 35-67). Boston, MA: Addison-Wesley.
- Scheuren, F. (2004). *What is a survey*. Alexandria, VA: American Statistical Association.
- Spencer, L., Ritchie, J., Lewis, J., & Dillion, L. (2003). *Quality in qualitative evaluation: A framework for assessing research evidence*. Retrieved April 22, 2006, from http://www.policyhub.gov.uk/docs/qqe_rep.pdf
- Stewart, T. R. (1987). The Delphi technique and judgmental forecasting. *Climatic Change*, 11, 97-113.
- Turoff, M., & Hiltz, S. R. (1996). Computer based Delphi processes. In M. Adler & E. Ziglio (Eds.), *Gazing into the oracle: The Delphi and its application to social policy and public health*. London, England: Jessica Kingsley.
- Uhl, N. P. (1983). Using the Delphi technique in institutional planning. In N. P. Uhl (Ed.), *Using research for strategic planning* (pp. 81-94). San Francisco, CA: Jossey-Bass.
- Wallsten, T. S., Budescu, D. V., Erev, I., & Diederich, A. (1997). Evaluating and combining subjective probability estimates. *Journal of Behavioral Decision Making*, 10, 243-268.
- Winkler, R. L., & Clemen, R. T. (2004). Multiple experts vs. multiple methods: Combining Correlation assessments. *Decision Analysis*, 1(3), 167-176.
- Winkler, R. L., & Poses, R. M. (1993). Evaluating and combining physicians' probabilities of survival in an intensive care unit. *Management Science*, 39(12), 1526-1543.
- Ziglio, E. (1996). The Delphi method and its contribution to decision-making. In M. Adler & E. Ziglio (Eds.), *Gazing into the oracle: The Delphi and its application to social policy and public health*. London, England: Jessica Kingsley.

Excerpt from Somerville, J. A. (2007). *Critical factors affecting the meaningful assessment of student learning outcomes: A Delphi study of the opinions of community college personnel*. Unpublished doctoral dissertation, Oregon State University, Corvallis, OR.