The Delphi Technique: Making Sense of Consensus

Chia-Chien Hsu
Brian A. Sandford

Follow this and additional works at: https://scholarworks.umass.edu/pare

Recommended Citation
Available at: https://scholarworks.umass.edu/pare/vol12/iss1/10

This Article is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Practical Assessment, Research, and Evaluation by an authorized editor of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
The Delphi technique is a widely used and accepted method for gathering data from respondents within their domain of expertise. The technique is designed as a group communication process which aims to achieve a convergence of opinion on a specific real-world issue. The Delphi process has been used in various fields of study such as program planning, needs assessment, policy determination, and resource utilization to develop a full range of alternatives, explore or expose underlying assumptions, as well as correlate judgments on a topic spanning a wide range of disciplines. The Delphi technique is well suited as a method for consensus-building by using a series of questionnaires delivered using multiple iterations to collect data from a panel of selected subjects. Subject selection, time frames for conducting and completing a study, the possibility of low response rates, and unintentionally guiding feedback from the respondent group are areas which should be considered when designing and implementing a Delphi study.

The Delphi technique, mainly developed by Dalkey and Helmer (1963) at the Rand Corporation in the 1950s, is a widely used and accepted method for achieving convergence of opinion concerning real-world knowledge solicited from experts within certain topic areas. Predicated on the rationale that, “two heads are better than one, or...n heads are better than one” (Dalkey, 1972, p. 15), the Delphi technique is designed as a group communication process that aims at conducting detailed examinations and discussions of a specific issue for the purpose of goal setting, policy investigation, or predicting the occurrence of future events (Ulschak, 1983; Turoff & Hiltz, 1996; Ludwig, 1997). Common surveys try to identify “what is,” whereas the Delphi technique attempts to address “what could/should be” (Miller, 2006).

In the literature, Delphi has been applied in various fields such as program planning, needs assessment, policy determination, and resource utilization. Delbecq, Van de Ven, and Gustafson (1975) specifically indicate that the Delphi technique can be used for achieving the following objectives:

1. To determine or develop a range of possible program alternatives;
2. To explore or expose underlying assumptions or information leading to different judgments;
3. To seek out information which may generate a consensus on the part of the respondent group;
4. To correlate informed judgments on a topic spanning a wide range of disciplines, and;
5. To educate the respondent group as to the diverse and interrelated aspects of the topic (p. 11).

CHARACTERISTICS OF THE DELPHI TECHNIQUE

The Delphi technique is well suited as a means and method for consensus-building by using a series of questionnaires to collect data from a panel of selected subjects (Dalkey & Helmer, 1963; Dalkey, 1969; Linstone & Turoff, 1975; Lindeman, 1981; Martino, 1983; Young & Jamieson, 2001). Delphi, in contrast to other data gathering and analysis techniques, employs multiple iterations designed to
develop a consensus of opinion concerning a specific topic. Ludwig (1994) indicates:

\textit{Iterations refer to the feedback process. The process was viewed as a series of rounds; in each round every participant worked through a questionnaire which was returned to the researcher who collected, edited, and returned to every participant a statement of the position of the whole group and the participant’s own position. A summation of comments made each participant aware of the range of opinions and the reasons underlying those opinions (p. 55).}

More specifically, the feedback process allows and encourages the selected Delphi participants to reassess their initial judgments about the information provided in previous iterations. Thus, in a Delphi study, the results of previous iterations regarding specific statements and/or items can change or be modified by individual panel members in later iterations based on their ability to review and assess the comments and feedback provided by the other Delphi panelists.

Other notable characteristics inherent with using the Delphi technique are the ability to provide anonymity to respondents, a controlled feedback process, and the suitability of a variety of statistical analysis techniques to interpret the data (Dalkey, 1972; Ludlow, 1975; Douglas, 1983). These characteristics are designed to offset the shortcomings of conventional means of pooling opinions obtained from a group interaction (i.e., influences of dominant individuals, noise, and group pressure for conformity) (Dalkey, 1972).

One of the primary characteristics and advantages of the Delphi process is subject anonymity which can reduce the effects of dominant individuals which often is a concern when using group-based processes used to collect and synthesize information (Dalkey, 1972). Additionally, the issue of confidentiality is facilitated by geographic dispersion of the subjects as well as the use of electronic communication such as e-mail to solicit and exchange information. As such, certain downsides associated with group dynamics such as manipulation or coercion to conform or adopt a certain viewpoint can be minimized (Helmer & Rescher, 1959; Oh, 1974; Adams, 2001).

Controlled feedback in the Delphi process is designed to reduce the effect of noise. Based upon Dalkey (1972), noise is that communication which occurs in a group process which both distorts the data and deals with group and/or individual interests rather than focusing on problem solving. As a result, the information developed from this kind of communication generally consists of bias not related to the purposes of the study. Basically, the controlled feedback process consists of a well organized summary of the prior iteration intentionally distributed to the subjects which allows each participant an opportunity to generate additional insights and more thoroughly clarify the information developed by previous iterations. Through the operation of multiple iterations, subjects are expected to become more problem-solving oriented, to offer their opinions more insightfully, and to minimize the effects of noise.

Finally, the ability to use statistical analysis techniques is a practice which further reduces the potential of group pressure for conformity (Dalkey, 1972). More specifically, statistical analysis can ensure that opinions generated by each subject of a Delphi study are well represented in the final iteration because, “at the end of the exercise there may still be a significant spread in individual opinions” (Dalkey, 1972, p. 21). That is, each subject would have no pressure, either real or perceived, to conform to another participant’s responses that may originate from obedience to social norms, customs, organizational culture, or standing within a profession. The tools of statistical analysis allow for an objective and impartial analysis and summarization of the collected data.

\section*{THE DELPHI PROCESS}

Theoretically, the Delphi process can be continuously iterated until consensus is determined to have been achieved. However, Cyphert and Gant (1971), Brooks (1979), Ludwig (1994, 1997), and Custer, Scarcella, and Stewart (1999) point out that three iterations are often sufficient to collect the needed information and to reach a consensus in most cases. The following discussion, however, provides guidelines for up to four iterations in order to assist those who decide to use the Delphi process as a data collection technique when it is determined that additional iterations beyond three are needed or valuable.

\subsection*{Round 1:} In the first round, the Delphi process traditionally begins with an open-ended questionnaire. The open-ended questionnaire serves as the cornerstone of soliciting specific information about a content area from the Delphi subjects (Custer, Scarcella, & Stewart, 1999). After receiving subjects’ responses, investigators need to convert the collected information into a well-structured questionnaire. This questionnaire is used as the survey instrument for the second round of data collection. It should be noted that it is both an acceptable and a common modification of the Delphi process format to use a structured questionnaire in Round 1 that is based upon an extensive review of the literature. Kerlinger (1973) noted that the use of a modified Delphi process is appropriate if basic information concerning the target issue is available and usable.

\subsection*{Round 2:} In the second round, each Delphi participant receives a second questionnaire and is asked to review the items summarized by the investigators based on the information provided in the first round. Accordingly, Delphi panelists may be required to rate or “rank-order
items to establish preliminary priorities among items. As a result of round two, areas of disagreement and agreement are identified" (Ludwig, 1994, p. 54-55). In some cases, Delphi panelists are asked to state the rationale concerning rating priorities among items (Jacobs, 1996). In this round, consensus begins forming and the actual outcomes can be presented among the participants’ responses (Jacobs, 1996).

**Round 3:** In the third round, each Delphi panelist receives a questionnaire that includes the items and ratings summarized by the investigators in the previous round and are asked to revise his/her judgments or “to specify the reasons for remaining outside the consensus” (Pfeiffer, 1968, p. 152). This round gives Delphi panelists an opportunity to make further clarifications of both the information and their judgments of the relative importance of the items. However, compared to the previous round, only a slight increase in the degree of consensus can be expected (Weaver, 1971; Dalkey & Rourke, 1972; Anglin, 1991; Jacobs, 1996).

**Round 4:** In the fourth and often final round, the list of remaining items, their ratings, minority opinions, and items achieving consensus are distributed to the panelists. This round provides a final opportunity for participants to revise their judgments. It should be remembered that the number of Delphi iterations depends largely on the degree of consensus sought by the investigators and can vary from three to five (Delbecq, Van de Ven, Gustafson, 1975; Ludwig, 1994).

**Subject Selection**

Regarding the selection of subjects for a Delphi study, choosing the appropriate subjects is the most important step in the entire process because it directly relates to the quality of the results generated (Judd, 1972; Taylor & Judd, 1989; Jacobs, 1996). Since the Delphi technique focuses on eliciting expert opinions over a short period of time, the selection of Delphi subjects is generally dependent upon the disciplinary areas of expertise required by the specific issue.

Regarding any set standards of selecting Delphi subjects, there is, in fact, no exact criterion currently listed in the literature concerning the selection of Delphi participants. That is, “throughout the Delphi literature, the definition of [Delphi subjects] has remained ambiguous” (Kaplan, 1971, p. 24). Regarding the criteria used to guide the selection of Delphi subjects, individuals are considered eligible to be invited to participate in a Delphi study if they have somewhat related backgrounds and experiences concerning the target issue, are capable of contributing helpful inputs, and are willing to revise their initial or previous judgments for the purpose of reaching or attaining consensus (Pill, 1971; Oh, 1974). Helmer and Rescher (1959), Klee (1972), and Oh (1974) concur that choosing individuals who are simply knowledgeable concerning the target issue is not sufficient nor recommended. Considering the necessity of selecting the most qualified individuals, Delbecq, Van de Ven, and Gustafson (1975) specifically state that three groups of people are well qualified to be subjects of a Delphi study. The authors recommend:

“1. the top management decision makers who will utilize the outcomes of the Delphi study;

2. the professional staff members together with their support team; and

3. the respondents to the Delphi questionnaire whose judgments are being sought” (p. 85).

Delphi subjects should be highly trained and competent within the specialized area of knowledge related to the target issue. Investigators need to closely examine and seriously consider the qualifications of Delphi subjects. Oh (1974) indicates that choosing appropriate subjects is generally based on the judgment and discretion of the principal investigators. Jones and Twiss (1978) state that the principal investigators of a Delphi study should identify and select the most appropriate individuals through a nomination process. Ludwig (1994) also states that, “solicitation of nominations of well-known and respected individuals from the members within the target groups of experts was recommended” (p. 52). Generally, the pool of selecting possible Delphi subjects is likely to use positional leaders (Kaplan, 1971; Ludwig, 1994), to follow a review of authors of publications in the literature (Meyer, 1992; Miller, 2001), and/or to make contacts with those who have firsthand relationships with a particular issue (Jones, 1975; Anderson & Schneider, 1993). The latter basically consists of individuals who are primary stakeholders with various interests related to the target issue or research effort.

Concerning the appropriate number of subjects to involve in a Delphi study, Delbecq, Van de Ven, and Gustafson (1975) recommend that researchers should use the minimally sufficient number of subjects and seek to verify the results through follow-up explorations. Ludwig (1994) notes that the number of experts used in a Delphi study is "generally determined by the number required to constitute a representative pooling of judgments and the information processing capability of the research team” (p. 52). However, what constitutes an optimal number of subjects in a Delphi study never reaches a consensus in the literature. Delbecq, Van de Ven, and Gustafson (1975) suggest that ten to fifteen subjects could be sufficient if the background of the Delphi subjects is homogeneous. In contrast, if various reference groups are involved in a Delphi study, more subjects are anticipated to
be needed. Witkin and Altschuld (1995) note that the approximate size of a Delphi panel is generally under 50, but more have been employed. Ludwig (1997) documents that, “the majority of Delphi studies have used between 15 and 20 respondents” (p. 2). In sum, the size of Delphi subjects is variable (Delbecq, Van de Ven, & Gustafson, 1975). If the sample size of a Delphi study is too small, these subjects may not be considered as having provided a representative pooling of judgments regarding the target issue. If the sample size is too large, the drawbacks inherent within the Delphi technique such as potentially low response rates and the obligation of large blocks of time by the respondents and the researcher(s) can be the result.

**Time Requirements**

Conducting a Delphi study can be time-consuming. Specifically, when the instrument of a Delphi study consists of a large number of statements, subjects will need to dedicate large blocks of time to complete the questionnaires. Delbecq, Van de Ven, and Gustafson (1975), Ulschak (1983), and Ludwig, (1994) recommend that a minimum of 45 days for the administration of a Delphi study is necessary. With regard to the time management between iterations, Delbecq et al. (1975) note that giving two weeks for Delphi subjects to respond to each round is encouraged.

Ludwig (1994) indicates, “a drawback to Delphi was that the questionnaire method may slow the process greatly as several days or weeks may pass between rounds” (p. 54). More specifically, since developing the instrument, collecting the data, and administering the questionnaire are interconnected between iterations, ensuring Delphi subjects respond to the investigators on time does in many ways either promote or prohibit the ability of the investigators in analyzing the data, developing a new instrument based upon the prior responses, and distributing subsequent questionnaires in a timely fashion. These are challenging aspects of conducting a Delphi study and do require proper planning and management.

The use and prevalence of electronic technologies (i.e., e-mail, teleconferencing, etc.) may facilitate those who are interested in using the Delphi technique. Witkin and Altschuld (1995) note that electronic technology provides an opportunity for individuals to more easily employ the Delphi process by taking advantages of, “(1) the storage, processing, and speed of transmission capabilities of computers; (2) the maintenance of respondent anonymity, and; (3) the potential for rapid feedback” (p. 204).

**Data Analysis**

Regarding data analysis, decision rules must be established to assemble and organize the judgments and insights provided by Delphi subjects. However, the kind and type of criteria to use to both define and determine consensus in a Delphi study is subject to interpretation. Basically, consensus on a topic can be decided if a certain percentage of the votes falls within a prescribed range (Miller, 2006). One criterion recommends that consensus is achieved by having 80 percent of subjects’ votes fall within two categories on a seven-point scale (Ulschak, 1983). Green (1982) suggests that at least 70 percent of Delphi subjects need to rate three or higher on a four point Likert-type scale and the median has to be at 3.25 or higher. Scheibe, Skutsch, and Schofer (1975) reveal that the use of percentage measures is inadequate. They suggest that a more reliable alternative is to measure the stability of subjects’ responses in successive iterations.

In the Delphi process, data analysis can involve both qualitative and quantitative data. Investigators need to deal with qualitative data if classic Delphi studies, which use open-ended questions to solicit subjects’ opinions, are conducted in the initial iteration. Subsequent iterations are to identify and hopefully achieve the desired level of consensus as well as any changes of judgments among panelists. The major statistics used in Delphi studies are measures of central tendency (means, median, and mode) and level of dispersion (standard deviation and inter-quartile range) in order to present information concerning the collective judgments of respondents (Hasson, Keeney, & McKenna, 2000). Generally, the uses of median and mode are favored. However, in some cases, as manifested by Murray and Jarman (1987), the mean is also workable. Witkin (1984) questions the appropriateness of using the mean to measure the subjects’ responses if scales used in Delphi studies are not delineated at equal intervals. In the literature, the use of median score, based on Likert-type scale, is strongly favored (Hill & Fowles, 1975; Eckman, 1983; Jacobs, 1996). As Jacobs (1996) states, “considering the anticipated consensus of opinion and the skewed expectation of responses as they were compiled, the median would inherently appear best suited to reflect the resultant convergence of opinion” (p. 57). The use of mode is also suitable when reporting data in the Delphi process. Ludwig (1994) specifically addressed that “the Delphi process has a tendency to create convergence, and though this was usually to a single point, there was the possibility of polarization or clustering of the results around two or more points. In these instances, the mean or median could be misleading” (p. 57).

**CONSIDERING DELPHI SHORTCOMINGS AND WEAKNESSES**

**Potential of Low Response Rates**

Due to the multiple feedback processes inherent and integral to the concept and use of the Delphi process, the
potential exists for low response rates and striving to maintain robust feedback can be a challenge. “In the Delphi technique, [poor response rate] is magnified fourfold because a maximum of four surveys may be sent to the same panelists” (Witkin & Altschuld, 1995, p. 196). If a certain portion of the subjects discontinue their responses during various stages of the Delphi process, the quality of information obtained could be discounted or at least critically scrutinized. As such, Ludwig (1994) specifically addresses subject motivation as the key to the successful implementation of a Delphi study and investigators need to play an active role in this area to help ensure as high a response rate as possible.

Consumption of Large Blocks of Time

The Delphi technique can also be time-consuming and laborious. Unlike other data collection techniques such as the telephone survey and the face-to-face administration, which can be simultaneously conducted by a group of people and can be completed in a short period of time if the sample size is small, the Delphi technique is iterative and sequential. As a result, the necessity of taking large block of time to successively complete a Delphi process is inescapable. Ludwig (1994) indicates that, “a drawback to Delphi was that the questionnaire method may slow the process greatly as several days or weeks may pass between rounds” (p. 54). Optimally speaking, the iteration characteristics of the Delphi process provide the opportunities for investigators and subjects to improve the accuracy of the results. In contrast, the same characteristic also increases the workload of investigators and the amount of time needed to successfully complete the data collection process (Cunliffe, 2002).

Potential of Molding Opinions

The iteration characteristics of the Delphi technique can potentially enable investigators to mold opinions (Altschuld, 2003). An experiment, conducted by Scheibe, Skutsch, and Schofer (1975), indicated that Delphi subjects would rate their responses differently after receiving a distorted feedback. Dalkey and Helmer (1963) also noted that, “some ‘leading’ by the experimenters inevitably resulted from the selection of the information supplied” (p. 467). Moreover, Cyphert and Gant (1971) illustrated that a statement in their study was initially rated below average. However, Delphi subjects rated the statement above average after receiving false feedback. Therefore, Cyphert and Gant (1971) concluded that the Delphi technique could, “be used to mold opinion as well as to collect [data]” (p. 273). Indeed, “subtle pressure to conform with group ratings” was one of the major drawbacks in the Delphi technique (Witkin & Altschuld, 1995, p. 188). Delphi investigators need to be cognizant, exercise caution, and implement the proper safeguards in dealing with this issue.

Potential of Identifying General Statements vs. Specific Topic Related Information

An assumption concerning Delphi participants is that they are equivalent in knowledge and experience (Altschuld & Thomas, 1991). However, this assumption might not be justified. More specifically, the expertise of Delphi panelists could be unevenly distributed, especially in the field of high technology (Marchant, 1988; Altschuld & Thomas, 1991). “Some panelists may have much more in-depth knowledge of certain topics, whereas other panelists are more knowledgeable about different topics” (Altschuld & Thomas, 1991, p. 187). Therefore, subjects who have less in-depth knowledge of certain topics are unable to specify the most important statements which have been identified by those subjects who possess in-depth knowledge concerning the target issue. The outcomes of a Delphi study could be the results of identifying a series of general statements rather than an in-depth exposition of the topic (Altschuld & Thomas, 1991).

SUMMARY

The Delphi technique provides those involved or interested in engaging in research, evaluation, fact-finding, issue exploration, or discovering what is actually known or not known about a specific topic a flexible and adaptable tool to gather and analyze the needed data. Subject selection and the time frames for conducting and completing a Delphi study are two areas which should be considered carefully prior to initiating the study. The additional precautions concerning low response rates, unintentionally guiding feedback, and surveying panelists about their limited knowledge of the topic rather than soliciting their expert judgments should also be built into the design and implementation of the study. The Delphi technique has and will continue to be an important data collection methodology with a wide variety of applications and uses for people who want to gather information from those who are immersed and imbedded in the topic of interest and can provide real-time and real-world knowledge.

REFERENCES


