

Using the Delphi technique in normative planning research: methodological design considerations

Nick Novakowski

Department of Geography/Environmental Studies, Memorial University of Newfoundland, Corner Brook, NL A2H 6P9, Canada; e-mail: nnovakowski@swgc.mun.ca

Barry Wellar

Department of Geography, University of Ottawa, Ottawa, ON K1N 6N5, Canada; e-mail: wellarb@uottawa.ca

Received 15 August 2006; in revised form 25 September 2006; published online 3 April 2008

Abstract. Worldwide, metropolitan areas continue to be confronted by a growing number of increasingly difficult planning issues. It is our experience that planning practitioners have not taken full advantage of what the Delphi technique can contribute to making informed choices in a wide variety of decision and policy environments. The objectives of this paper are to describe and explain the research design that supported a real-world application of the Delphi technique in an urban, regional, and ecosystem-based planning context, as well as to demonstrate how this model has been or can be adapted to serve a variety of planning research or application tasks.

Introduction

Delphi is the site where the most revered oracles of ancient Greece formulated their predictions about the future (De Boer and Hale, 2000). In contrast to the predictions provided by the Pythia or priestesses, however, the research results derived from use of the Delphi technique are driven by methodological design. While instances of the use of the Delphi technique are evident in many disciplines, details regarding the employment of the technique in the planning literature have been scant since the 1980s. To address this void, we briefly outline the steps of the Delphi technique, and present methodological design findings from a real-world application of the technique used to derive the normative characteristics of high-quality plans generated within the framework of ecosystem-based planning. The paper concludes with lessons learned from that real-world exercise, which will be the subject of a companion paper.

The Delphi technique: background

The Delphi technique was developed by Dalkey and Helmer at the Rand Corporation in 1953 to explore the potential bombing strategies that Soviet military leaders might implement in the event of an atomic war (Dalkey, 1969; Helmer, 1983; Linstone and Turoff, 1975a; Rowe and Wright, 1999). The client of the Rand Corporation was the US Air Force, which recognized the complexity of the subject matter and needed a way to utilize the considerable expert knowledge on strategic bombing that existed in America after World War II.

Once the military applications of the Delphi technique were declassified in the 1950s and it was first publicly described in 1964, the technique entered the academic mainstream (Martino, 1999).

“Delphi may be characterized as a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem” (Linstone and Turoff, 1975a, page 3).

In planning, the individuals involved in the group communication are typically planning experts. The structure of the group communication is designed by a *monitor* or *monitor team* that formulates a reiterative survey to address the research topic. The survey is sent to the designated group of experts (known as the *Delphi panel*) who then anonymously rank their preferences regarding a continuum of answers related to a series of questions or propositions posed. The experts subsequently return their responses to the Delphi monitor. Each of the iterative mail-outs (either by conventional mail or by e-mail) of the survey is called a *round*, and rounds continue until stable responses between rounds are achieved.

After the first round the monitor reviews and summarizes the responses, and then employs a measure of central tendency (usually the mean, median, or mode) to indicate where the majority of the panel responses are located on the response continuum. The response continuum can be based on Likert scale categories of five to seven points, or on some other rationale, in order to indicate the degree to which panel members agree or disagree with the questions or propositions posed (Critcher and Gladstone, 1998).

The monitor then develops a second-round survey that reveals the response dispersion of the panel, and also includes the feedback obtained from the first round's open-ended question(s) to ensure that the survey author has not overlooked something relevant to the topic. Upon receipt of the second round, the experts are asked to consider the position of the measure of central tendency of the panel, and are permitted to revise their initial responses if they choose. Although the process of response and reiteration can be repeated as many times as required, Delphi practice has revealed that the rate of response convergence is highest between rounds 1 and 2 (Linstone and Turoff, 1975a).

Although there are a number of different types of Delphi exercises, they can usually be assigned to one of three broad categories:

1. *Normative Delphi*. Normative Delphi exercises are explorations of *what should be*, given current knowledge (Martino, 1999). Obtaining consensus about a preferred future state or process is typically the primary research objective. An example of a normative Delphi includes the research of the authors, where a ten-member expert planning panel was employed to generate and measure agreement about the characteristics of high-quality ecosystem-based plans (Novakowski, 1999). The Delphi technique was employed to derive an evaluation framework that was used to ascertain which plan form, plan content, and planning process conditions are necessary and sufficient for the evaluation of ecosystem-based plan quality. The results of the Delphi exercise were then subjected to metacriteria analysis and interpreted by employing hierarchy theory. The research results are presented in a forthcoming paper.

2. *Forecasting Delphi*. A forecasting Delphi exercise is concerned with predictions about future events for which little, diverse, or conflicting knowledge currently exists (Albright, 2002; Coates et al, 2001; Ilbery et al, 2004). While there are four primary approaches to forecasting (extrapolation, leading indicators, causal models, and stochastic methods), the Delphi panel may intuitively use one or all four of the approaches (Martino, 1999). A relatively recent forecasting Delphi investigated whether feedback and discussion within the exercise would improve the accuracy of either individual or group predictions regarding volatile world events (Parenté et al, 2005).

3. *Policy Delphi*. The policy Delphi involves the exploration of a matter of political interest or consequence (Coates, 1999; Critcher and Gladstone, 1998; Turoff, 1975). The purpose of the policy Delphi is not expert consensus; rather, it is concerned with ensuring that the range of politically relevant variables and contextual parameters is identified and explored (Wellar, 1997). A real-world example of the policy Delphi involves Canada's Flood Damage Reduction Program, where a fifty-member panel

explored the benefits and costs of the program in terms of the improved management of hazard areas, protection of significant environmental features, higher costs for developers, and other policy elements (De Loë and Wojtanowski, 2001).

As a research process, each Delphi category shares the same overall structure in terms of group communication, anonymity, iteration, and central tendency. However, due to differences in research objectives and/or application domains (identification of preferred futures, forecasting, or policy alternatives), a choice must be made about which type of Delphi exercise is appropriate for a particular planning research issue or problem.

Relevance of the Delphi technique to urban and regional planning research

The Delphi technique demonstrates utility and applicability to understanding, representing, and adjudicating planning issues. Linstone and Turoff (1975a, page 4) state that the Delphi technique is particularly useful when the research problem “does not lend itself to precise analytical techniques but can benefit from subjective judgments on a collective basis.” Richey et al (1985, page 142) observe that the Delphi technique “lends itself more readily to conceptual or philosophical issues than to issues that require exact, or quantitative, answers.” Similarly, as Ying and Kung (2000) suggest, use of the Delphi is suitable when objective observation of data is neither feasible nor possible, a view that is corroborated by Critcher and Gladstone (1998). In principle, then, it appears fair to state that the Delphi technique is highly pertinent to different types of planning research.

The Delphi technique also provides an equal opportunity for participants to express their opinions since responses are neither ranked nor weighted according to participant. In other words, all panelists contribute equally to the measure of central tendency since different voices are heard and different opinions and value systems are ‘put on the table’ for consideration by participants. This feature of the Delphi technique is consistent with the tenets of participatory planning, public participation in planning, and the use of advisory groups in planning, which similarly seek to ensure full and fair involvement of all parties with an interest in a planning action.

Reliance on expert opinion has characterized decades of planning practice in metropolitan areas worldwide. At the same time, there is a range of potential problems associated with eliciting expert opinion in consensus-building exercises: the influence of dominant personalities; the reluctance of experts to change their minds even in the presence of contradictory facts; and the psychological desire, even among experts, to conform to majority opinion (Martino, 1999). As a result, it is emphasized that careful attention needs to be paid to designing a research question or proposition in such a way that it elicits the very best thinking that experts can bring to bear.

Transparency and the design of a Delphi exercise

In this section, a practical research design for the Delphi technique that was tested in a real-world planning context is presented. The Delphi type undertaken was a normative Delphi, and was employed to identify and test the relative importance of the characteristics of high-quality municipal plans formulated in an ecosystem-based planning context. Stages in the proposed process are described and explained in a step-by-step fashion (figure 1).

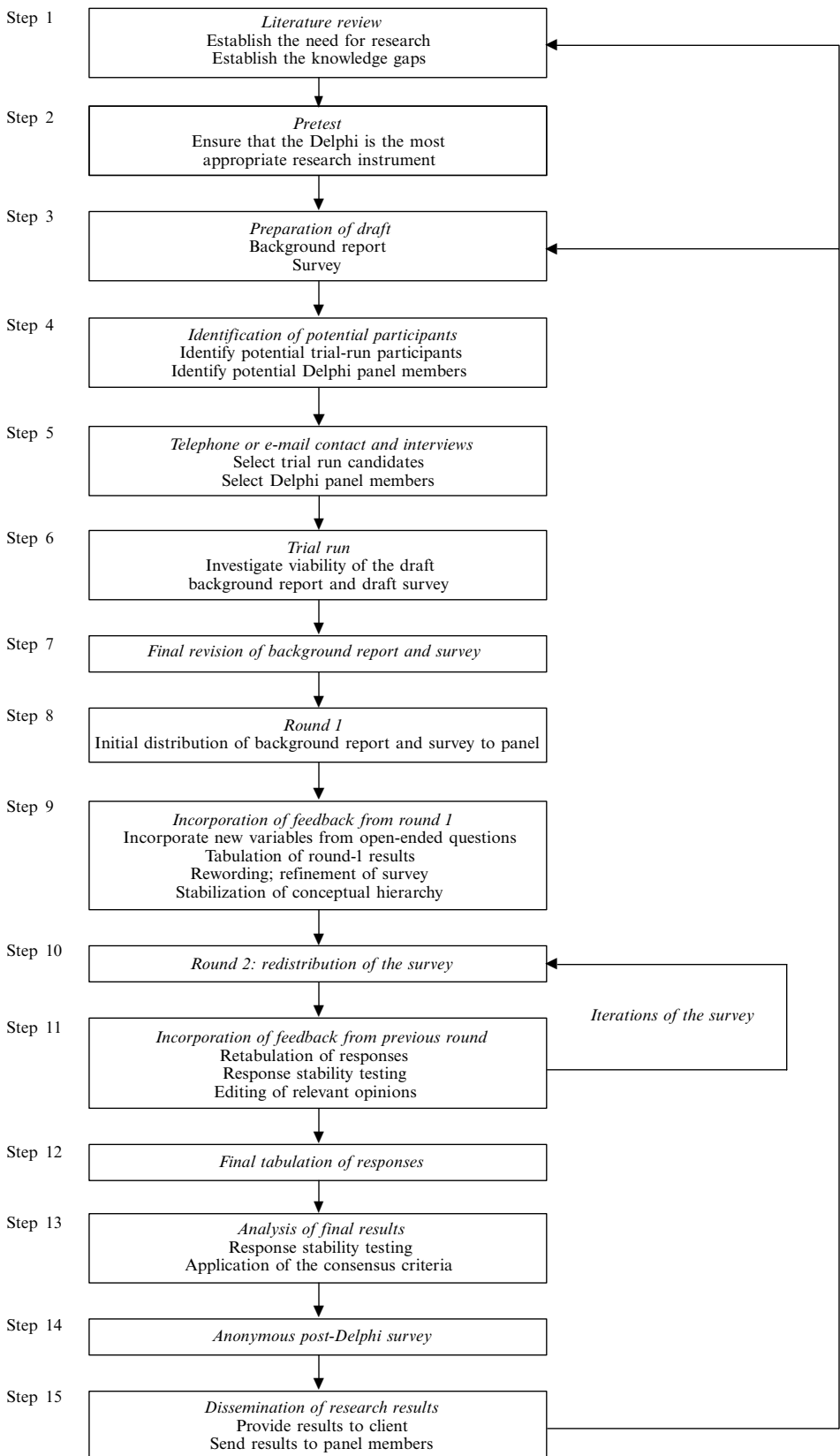


Figure 1. Flowchart for a normative Delphi.

Step 1: literature review

A systematic literature review is the point of departure for any Delphi exercise. The literature review serves such critical functions as the following: (a) preparing a comprehensive statement of current knowledge concerning the subject matter (that is, what is known); (b) identifying subject matter knowledge gaps, dominant research calls, theoretical shortcomings, or data-based inconsistencies (that is, what needs to be known); and (c) providing insight into potential experts to involve in the Delphi as advisors and/or participants. Identification of knowledge gaps is crucial to the success of the Delphi since those findings establish whether the research design is exploratory or confirmatory in nature. Further, the literature review enables the researcher or Delphi monitor to give the initial survey shape and content, and contributes to an understanding of whether the Delphi technique is the appropriate method for the study.

Due to the critical importance of the literature review, it is imperative that the search procedure upon which the review is based be as robust as time and other resources allow. The following kinds of literature have been identified as being among those that constitute what is collectively known as ‘the literature’: learned; popular; corporate/institutional—public; corporate/institutional—private; legal; regulatory; professional group; public interest group; and the vested/special interest group (Wellar, 2005). Clearly, since errors of both omission and commission can have major impacts on study resources and credibility, great care must be taken when making the decision about which bodies of literature to include or exclude in this all-important first step of a Delphi project.

Step 2: the pretest

Although the terms ‘pretest’ and ‘trial run’ are sometimes used interchangeably in the research design literature, Ackoff (1953) provides an explicit differentiation of these two research procedures. According to Ackoff (1953, pages 336–337), the pretest provides the opportunity to investigate alternative operational procedures for undertaking the research. The trial run subsequently examines the viability and efficiency of the specific instrument and the research plan before the full survey is launched, and evaluates whether the research plan and instrument—as designed—are likely to be effective and efficient. Unlike the pretest that is concerned with the comparative consideration of alternative techniques or methods, the trial run is directed towards an overall research plan (Ackoff, 1953, page 344).

In the absence of real-world/empirical evidence to direct the choice of a study procedure, engaging experts in a project can be considered an acceptable place for the researcher to begin (Wilhelm, 2001). If we work from the assumption that an expert-based technique is the appropriate way to proceed, the central task of the pretest involves selecting which technique involving expert opinion is the most suitable. There are many techniques that can be used to elicit expert opinion other than the Delphi, including surveys, the panel evaluation method, professional polling, the committee approach, roundtables, cross-impact analysis, brainstorming, or workshops. Each research technique has advantages and disadvantages that need to be considered during the pretest.

Step 3: preparation of the draft background report and survey

If it is decided that the Delphi technique is the appropriate research procedure, then the next step is to draft a background report for the panel explaining the *why*, *what*, and *how* features of the inquiry. Before panel selection, it is important to have the draft background report and survey prepared so that the specifics of who needs to be selected or what interests need to be represented are clear. Further, providing the panel

with a diagram that details how the stages of the process are intended to unfold was substantiated as very helpful in our Delphi exercise (see figure 1). The draft survey (which contains the actual propositions or statements being investigated) accompanying the draft background report can later be treated as either a full-blown round or as an information-seeking round called a round 0 (Helmer, 1983). A round 0 is employed to expose the Delphi subject matter to expert brainstorming and discourse. After panel selection in the next step, a decision about whether or not to use a round 0 is required.

Due to the complexity of some planning issues and the large number of details that can be involved in each round, a round 0 (which does not count as a formal round since responses are not counted or tabulated) can constitute an extra time imposition on panel members and therefore contribute to panel fatigue. This is a particularly important consideration if panelists are not being paid. If the exercise is straightforward, and/or the panelists are being paid, then a round 0 is in order. Conversely, if the exercise is likely to be time-consuming and arduous, then a round 0 should likely be eliminated. Alternative means of achieving the outcome of a round 1 without engaging the panel include the use of expert interviews, focus groups, walk-through and think-aloud exercises, and directed or targeted keyword-based literature searches that contribute to expanding the survey instrument to include the full envelope of policy items, problem solutions, or design features that need to be explored (Turoff, personal communication, 28 August, 2006; Wellar and Vandermeulen, 2000).

Meanwhile, the wording of the survey questions or propositions follows standard procedural guidelines for survey design. Variables are presented in plain language with minimal complexity so that ambiguity is minimized. Brevity is important in order to optimize clarity and speed. On the one hand, if the wording of the questions or propositions is too concise, then excessive freestyle interpretation by the panelists can result. On the other hand, if the questions or propositions are too lengthy, they may require the assimilation and/or interpretation of too many dimensions, and this can lead to confusion (Linstone and Turoff, 1975b, page 232). This concern is particularly relevant in two situations: when selecting an international panel, as cultural uses of professional terms and the idiomatic use of language may differ; and when designing a Delphi exercise that involves participants with very different backgrounds, such as elected officials, professional staff, and 'regular citizens' (Wellar and Vandermeulen, 2000).

The challenge, therefore, is finding the compromise between all that could be said and that which needs to be said to ensure that participants have a shared understanding of the core meaning of the variables involved. Preparation of the draft background report and initial survey brings the researcher to the point where state-of-the-art issues have been identified and consideration of panel member selection can begin. Simultaneously, since a heterogeneous panel is to be selected, the removal of ambiguity in the draft background report and survey is crucial as the Delphi team will perceive the subject matter from various perspectives.

Step 4: identification of potential participants in the Delphi panel

The first question concerns the size of the panel, which in turn will affect the stringency of the selection criteria. Panel size can vary widely. In a general sense, Turoff (personal communication, 1 June, 2006) writes that,

"The use of from three to five experts usually resulted in overlapping explanations. Therefore most of us practising the technique take the topic and ask how many different types of experts do we need to examine it from all relevant perspectives? Multiply this by five and you have the total number that should be in the panel and after you invite them if you have at least three in each category that have agreed you might go with that."

A critical consideration arising from Turoff's observation is the establishment of "all the relevant perspectives", which in turn indicates the scope of the Delphi and the size of the panel. More specifically, for planning-related issues or missions, the literature suggests that a "typical Delphi panel has about 8 to 12 members" (Cavalli-Sforza and Ortolano, 1984, page 325). A similar guideline is suggested by Richey et al (1985, page 142), who state that "a small panel (e.g., eight) would be sufficient to develop appropriate consensus views."

As demonstrated by Wellar's work on the Walking Security Index project, however, the numbers for a Delphi transportation planning project can justifiably reach well into double figures. Based on his client-driven study, the mix of potential panelists could include elected officials, community association leaders, an array of citizens (seniors, adults, youth, children, pedestrians, cyclists, and so on), and professionals from planning, engineering, public health, and law enforcement (Wellar, 1997; Wellar and Vandermeulen, 2000). The research undertaken by Wellar underscores Turoff's suggestion that the first priority is to ensure that all the perspectives on the issue are addressed through panel composition; the second priority is panel size.

The actual identification of potential participants for the trial run and final run requires consideration of the term 'expert', viz what are the attributes that characterize an expert? The legal aspects of the term 'expert' can be employed as a starting point. The legal view of 'expert' and what it is constituted by represents 'the stock and trade' of expert witnesses used in planning-related jurisprudence. In a legal context, opinion-based evidence (as opposed to observation-based evidence) is normally admissible in court only when it is provided by an expert. We hasten to add, however, that in a policy Delphi 'ordinary citizens'—that is, voters—could be regarded as experts on municipal issues. The point being emphasized is that there is room and a need to be open-minded about the knowledge or expertise pertinent to the Delphi exercise under consideration.

With the preceding remarks as context, the following criteria can be used to guide the selection process for the expert panel: an advanced degree in disciplines related to the research domain; a relevant publication record demonstrating professional or academic interest; extensive related work experience in the research domain; professional affiliation (eg the Canadian Institute of Planners, the Royal Town Planning Institute, the American Planning Association); and gender, ethnicity, life-cycle stage, or other factors to the extent that these characteristics are relevant a priori to the research topic. Overall, Delphi monitors "seek to create a panel that reflects a wide range of experience and a diversity of opinions on the subjects that are being considered" (Masser and Foley, 1987, page 218).

At the same time, knowledgeable people rather than subject matter or methods/techniques experts are also pertinent to normative preference probes. Planning issues affect and are affected by the people living within the purview of planning decisions (eg downwind from a new toxic waste incinerator, or park users facing the loss of local green space). As Ziglio (1996, page 14) remarks, "the definition of 'experts' varies according to the context and field of interest." As such, expertise can be both lay and professional in a normative preference probe.

Step 5: telephone and e-mail contact and interviews

The original list of potential participants can be derived from the literature review, from profiles of community activity, from membership directories, and/or from the public record. Potential panel members can be contacted either by telephone or by e-mail, starting with potential participants who appear to have made a significant contribution to one or more of the selection criteria. The monitor can also ask each contact to suggest other suitable participants, a practice known as *snowballing*. Assurances about panel anonymity need to be provided and stressed in this stage.

Step 6: the trial run

Since the Delphi technique has been decided upon in the pretest, a first consideration in the trial run concerns whether the emerging web-based Delphis are of interest to the survey designers. One new incarnation of the Delphi process involves permitting the panelists to change their response at any time during an online exercise. New computer-based versions allow respondents “to participate in any phase at any time and eliminates sequential constraints for each individual” (Turoff et al, 2004). The use of computer-mediated communication is a new and promising direction for Delphis involving larger panels. Turoff and Hiltz (1996) make the observation that rounds might even be eliminated in computerized Delphis by making the process continuous where individuals could respond to different phases at different times, and others would be notified automatically when a new addition to the discussion was made.

Linstone and Turoff (2002) make the primary distinction among types of Delphi as being between the ‘conventional Delphi’ (discussed here) and the ‘Delphi Conference’, which uses computer technologies to replace the Delphi monitor and compile the survey results through algorithms. “This latter approach has the advantage of eliminating the delay caused in summarizing each round of Delphi, thereby turning the process into a real-time communications system” (Linstone and Turoff, 2002). Regardless of whether a conventional or conference Delphi is selected, robust methodological design in the interest of generating replicable results remains the goal.

The trial run is undertaken to determine whether any final adjustments to the background package and the survey instrument itself are required before the final run. The trial run should involve participants who meet as many of the panel-selection criteria as possible, but have not been targeted to be on the final panel. Typically, the trial run will reveal ambiguities where further information needs to be provided, and whether there are variables or subject attributes that have been overlooked. The trial run is critical to ensuring that the survey instrument and its background report are both efficient and effective. In our Delphi trial run, we polled respondents who met all of the selection criteria except the one for publishing.

Step 7: final revision of the background report and survey

After completion of the trial run, the background report and survey can be fine-tuned and prepared for round 1. At this stage it is conceivable that questions could arise about suspending the anonymity rule. It is our experience (for reasons given below) that anonymity is best regarded as a required condition throughout the process, including the shift from the trial run to round 1.

Step 8: round 1

Once the final version of the background report and the survey are ready, the next step is to distribute them to the panel members. Panelists can be contacted by telephone or e-mail to alert them to the onset of round 1. Several days after the send-out, the monitor can contact panelists to ensure that materials have been received, and to answer any questions that arise. The monitor needs to respond quickly, unambiguously, and consistently to requests for information. An important element of round 1 is the inclusion of an open-ended question (or questions) which is intended to elicit feedback on any important relationship or variable that may have been overlooked, and which can later be incorporated into round 2.

In round 1 the Delphi monitor may choose to improve the level of response accuracy by asking panel members to rate their confidence in their answers. Turoff (personal communication, 1 June, 2006) suggests that “if you ask the respondents to rate their confidence in the answer and weigh the estimates by the confidence rating the result is better than an average of all the estimates.” Another approach used by

Turoff is to include a 'no judgment' choice among the options when panelists do not seem ready to address a particular variable. Turoff (personal communication, 28 August, 2006) finds this to be an excellent way to deal with highly heterogeneous panels that have wide-ranging areas of expertise and experience. In a related vein, one of us (Wellar) has been a member of two national funding panels in the past year in which in a rating system was used to ascertain the ability of experts on the panel to assess research proposals, and to then assign first, second, and third reader responsibilities for the proposals accordingly. This is a variation on the Turoff experience, and in combination they illustrate an important avenue available to monitors wanting to increase the validity of responses.

Step 9: incorporation of feedback from round 1

When all panel members have responded to round 1 and have returned their results to the Delphi monitor, the survey is revised in three primary ways:

- (1) a measure of central tendency is provided to identify the dominant response category along the response continuum for each statement or proposition in the survey;
- (2) new questions, propositions, or variables that have been suggested by the panelists in response to the open-ended question(s) are added. Not only does this demonstrate to panelists that their feedback has a substantive impact on the process, but also it permits the panelists to test their own ideas and hypotheses, and thereby increases their sense of ownership in the process; and,
- (3) individual questions and variables can be refined (rephrased, reworded) as a result of suggestions from the panel, but without altering the core meaning of the variable being tested.

Since planning-based Delphi exercises are often normative preference probes, a useful measure of central tendency to employ is the *mode*. Since the mode is located in the response category (in the distribution) with the largest number of observations, it can be regarded as the preferred parameter/statistic for identifying the panel's position on the response continuum.

If the distribution of votes along the continuum of responses is evenly dispersed, this may indicate a situation where there is a need for clarification or additional information to help respondents to reach consensus (Turoff, personal communication, 28 August, 2006). Another situation could be panel polarization at either end of the response continuum. In the latter case, soliciting additional information on the variable from the panel members is needed in order to gain insight into the polarization.

Step 10: redistribution of the survey

Round 2 is sent as soon as the survey has been revised and the centrality measures have been calculated. If a response more than one interval away from the modal response is to be chosen in round 2, then panelists can be asked to specify why. This approach worked well in our municipal plan quality Delphi, and helped to express the full array of perspectives involved in the discussion. Further, it promoted gaining insight into residual dissensus by (a) exploring the differences of opinion and (b) identifying factors that were considered by some experts but not by others (Widstrand and Kruus, 1996, page 61). As we discovered, dissensus can arise due to differences in perspectives, knowledge bases, interpretations of variables, theoretical views, and/or disciplinary bias (Helmer, 1983, page 134), and regard for different factors to take into consideration can arise for similar reasons.

In round 2 the members of the Delphi panel can be asked to reconsider the variables in light of the identified measure of central tendency or anchor (in our exercise, the mode was employed). Normally, there are three ways in which panelists respond to the measure of central tendency:

-
- (1) They ignore it and hold to their original selection.
 - (2) They shift judgment away from the mode by responding in an extreme fashion in the interest of moving the centrality anchor closer to their true desire. For example, an ongoing concern with Delphi results is that individual panel members may intentionally bias or overaccentuate their responses in order to consolidate their own research or another agenda (Jones, 1975, page 160). This is an important consideration if the Delphi results can have implications for funding or professional development.
 - (3) They shift judgment towards the mode, which is known as *assimilation* or *convergence*.

The Delphi monitor has the option of sending a copy of a panelist's round-1 responses along with round 2. If the panelists are not being paid, the Delphi monitor may decide that a further imposition on panelist effort levels is unsuitable and can therefore refrain from sending the round-1 responses.

Step 11: incorporation of feedback from previous round

Feedback from round 2 includes reasons why panel members are choosing non-modal responses, and this information can be provided in an inventory with round 3. Without exception, all feedback of this nature is to be formally recorded and transmitted back to the panel. Grammatical standardization (eg verb tense) should be the only editing done to panel comments. Generally, three rounds are sufficient to attain stability in Delphi responses: "Further rounds tended to show very little change and excessive repetition was unacceptable to participants" (Linstone and Turoff, 1975b, page 229).

Step 12: final tabulation of responses

Delphi responses to individual propositions and/or statements within the survey should be considered final only when they are stable. Many PhD dissertations continue to employ levels of panel consensus as *the stopping criterion*, despite widespread recommendations in the Delphi literature suggesting that consensus criteria should be applied only after response stability has been established (see Chaffin and Talley, 1980; Dajani et al, 1979; Regier, 1986; Sharma and Gupta, 1993). As Schiebe et al (1975, page 262) observe, using response stability rather than consensus levels as the stopping criterion "allows much more information to be derived from the Delphi." Once the responses are deemed stable, then the level of consensus criteria can be applied.

Stability itself "refers to the consistency of responses between successive rounds of a Delphi survey" (Chaffin and Talley, 1980, page 67). In other words, stability is a measure of the extent and degree to which panel members are selecting the same responses between successive rounds. One important distinction needs to be addressed: the stability of individual responses suggests the stability of the group, but not vice versa (Chaffin and Talley, 1980, page 70). This means that individual response stability represents a more stringent test. Following the suggestion of Regier (1986), we suggest that individual response stability be used both because the logic of doing so is sounder, and because the testing process can be made efficient. In fact, Nelson (1978, page 45) provides a practical rule of thumb and asserts that variables can be "considered stable when 20% or less of the participants changed their responses." In other words, if two or fewer of a hypothetical ten panelists are changing their responses concerning the importance rating of individual propositions or statements between rounds, then the individual responses can be considered stable. As soon as all of the variables in the survey are identified as stable, then the level-of-consensus criteria can be imposed.

Step 13: treatment of the final results of the Delphi exercise

As soon as the responses in the survey are identified as stable, the survey iteration process is halted and the *level-of-consensus* criteria are applied. To build on the work of Dajani et al (1979) by adding the new level-of-consensus criterion of *bimodality*, six possible agreement levels or stopping criteria can be recognized: consensus, majority, bimodality, bipolarity, plurality, and disagreement. The level-of-consensus criteria are defined as follows:

- Complete consensus: occurs when the response category has unanimous agreement among all respondents.
- Majority: occurs when more than 50% of the respondents choose the same response category. Once a stable response has been demonstrated, *majority* can be used as a decision rule (which means that the responses can then be used as a research result).
- Bimodality: occurs when an equal number of respondents are grouped in two response categories. When bimodality results, the precautionary principle of ecology can be imposed where the more stringent or safe response can be selected from the two alternatives.
- Bipolarity: occurs when respondents are consolidating in two response categories at opposite ends of the response continuum.
- Plurality: occurs when a large portion of the panel (but less than 50%) reaches stability in different response categories.
- Disagreement: occurs when respondents maintain views independent of other respondents across a range of response intervals.

It warrants noting that variables demonstrating stability after round 2 can be removed from subsequent iterations. This action is capable of contracting subsequent rounds—perhaps substantially (approximately 30% in our case)—and represents an important design decision for an efficient Delphi exercise. Once the results of the Delphi are itemized through consensus analysis, then the data can be treated or interpreted in various ways, including metacriteria analysis, theoretical consistency testing, operational feasibility testing on a case study, or by inspection within an appropriate theoretical and/or conceptual framework. In our case, we took the results of the Delphi and interpreted them within an ecological theory (hierarchy theory).

Step 14: anonymous post-Delphi survey

A post-Delphi survey is undertaken for two primary reasons: to gain insight into (1) how the process might be improved; and (2) how reliable the panelists feel that their responses are. In addition, however, the post-Delphi survey provides panel members with a sense of ownership in the project. The normative Delphi, for example, is a consensus-building exercise, and participants need to be assured that their commendations or condemnations of the process are appreciated. More generally, the post-Delphi survey can provide insight into both the effectiveness and the efficiency or delivery of the process.

Step 15: dissemination of research results

The dissemination of research results constitutes a key building block dimension of science. Several types of dissemination are available such as: using the results to shape plan, policy, or program content and delivery; publishing the results in one or more of the relevant literatures; and providing the monitor and panel members with a summary of research results for circulation among members of their networks.

Methodology-related research results from our specific application of the Delphi process

According to Masser and Foley (1987, page 217), the Delphi is “the most widely used technique for eliciting expert opinion.” However, recent (July 2006) Internet searches show that there is proliferating confusion about the differences between the Delphi and other consensus-building techniques, and in particular the methodological design aspect. In the interest of improving the methodological design considerations that underlie and direct Delphi exercises, and contributing to a better understanding of how Delphi differs from other associated techniques, we next present a selection of methodological lessons learned from a completed Delphi application involving the evaluation criteria of municipal plans.

Some of the potential pitfalls and obstacles to be overcome when implementing the Delphi technique have been discussed previously (Linstone, 1975; Linstone and Turoff, 2002; Regier, 1986). Additional lessons learned that arose from the Delphi application underlying this paper include the following:

1. *Differentiation between the pretest and the trial run is important.* The pretest provided the opportunity to elicit which expert-based technique would work best and actually yield the sort of results that we needed. In the pretest, cost considerations were paramount and this precluded any technique that involved convening the experts at a specific location. A smaller panel was desired, so a conventional rather than conference Delphi was selected. At the same time, it is important to remember that Delphi exercises are time consuming and tend to take longer to execute than other expert-based techniques.
2. *Doing a trial run.* The trial run actually takes a full-blown draft Delphi survey and tests it on a sample panel (as a test run to fine-tune the survey). In our case we used a proxy panel that met all of the selection criteria except one: a publishing record on the topic. Consequently, we engaged an entirely different set of highly qualified people for the trial run.
3. *Panel size.* We concur with Cavalli-Sforza and Ortolano (1984) that using a panel size of eight to twelve may be appropriate in many cases. The panel size of ten that we used in our municipal plan evaluation Delphi worked well, provided a diversity of expert opinion, and permitted us to engage the very best North American experts on the topic without being overly onerous for a single Delphi monitor. In retrospect, however, it is apparent that using an odd number of panelists (eg nine or eleven) would have eliminated the possibility of bimodal responses and therefore made interpretation more concise. In a different project, however, panel size was a secondary consideration, with the driving priority being that of ensuring that all perspectives that warranted consideration were represented.
4. *Strict panel selection criteria.* As stated, the panel selection criteria that we employed were strict. Using strict selection criteria not only served to provide high-quality responses, but also meant that the panelists felt validated by the experience after the (unexpected but later agreed upon) ‘reveal’ regarding panel identities.
5. *Design to prevent intentional response polarization.* As mentioned, reactions in round 2 to the mode by panelists can result in responses that move away from the mode in an extreme fashion in order to displace the dispersion, make a point, and/or bring the response mode closer to the panelist’s true view. This possibility can be reduced by providing fewer response categories than the traditional Likert scale using five or seven options.
6. *Maintain panel anonymity until the post-Delphi survey is completed.* During the execution of the process, it was our experience that there was much guessing among panelists regarding the identity of the other panelists. It is vital not to engage with any panelists on this matter. Further, we believe it is preferable that panelists are not told that there would be an option to reveal their identities once the process was completed.

7. *Difficult personalities.* We addressed this matter during snowballing since difficult personalities can seriously undermine the research panel. We found that personal comments about other potential panel members were common, but the personality factor did not materialize in our Delphi.

8. *A potential dark side to the Delphi process: angry panels.* Even if the Delphi monitor has attempted to minimize the participation of difficult personalities during the panel selection process, angry panels can still result with normative preference probes because value systems are involved. The job of the Delphi monitor is to remain impartial and to communicate the explanatory comments of panelists with a minimum of editing. Nevertheless, editing should be imposed to remove derogatory or strongly emotional material in the feedback between rounds. There are exceptions in the latter case, however, which could include Delphi exercises where a political entity is the sponsor (Turoff, personal communication, 28 August, 2006).

9. *Revealing the identity of the panel.* The anonymity of panel members can be maintained in perpetuity. Alternatively, once the post-Delphi survey has been completed, the identity of panelists can be revealed upon receipt of approval to do so. If the panelists are happy with how the Delphi exercise was executed, then they will be much more likely to reveal their own identities. And, if the panel has been carefully selected and true leaders in the field have been employed, then this will provide further validation that they have participated in a worthwhile study. An exception to this practice would be when a difficult interpersonal situation arises and it becomes necessary to maintain everybody's participation by invoking the reputations of the other participants (Turoff, personal communication, 28 August, 2006).

10. *Compensation.* The Delphi technique involves the use of experts, and the time of experts is valuable. Ideally, compensation for the panelists should be involved, particularly if the study will take more than 3–4 hours of their time. Alternatively, for some panelists, the opportunity to be engaged in a well-organized and well-executed Delphi exercise can be sufficient reason to participate. For others, compensation should take the form of either financial or research support for the panelists. For students doing research, they should recognize that panelists have their own research interests and may appreciate research support from a researcher who is familiar with their area of expertise. This is a quid pro quo solution and one of the options employed by the authors of this study. Without exception, formal letters of thanks sent to the employers, supervisors, or professional associations of participating panelists should be provided.

11. *Dealing with panel fatigue.* So-called 'panel fatigue' is a true concern. Researchers need to make every effort to minimize the size and complexity of the questionnaire and use e-mail. E-mail response rates are much quicker, an observation that was easily demonstrated by a mid-course mail-out in our survey process. In fact, the majority of the panel stated that turnaround time would be considerably quicker due to the overall immediacy of e-mail. However, surveys of considerable complexity generated in a computer application that is not widely used may represent an obstacle to the use of e-mail.

12. *Dealing with varying response rates among panel members.* The salient point to remember is that experts are busy people and their turnaround time is being juggled with an extensive range of commitments. Although the use of e-mail or couriers may expedite the process somewhat, the real issue is the prioritization of the research effort by the experts involved. Without providing financial compensation, the only tool that the Delphi monitor can employ is moral suasion. The salient point is the following: researchers must allow themselves enough time for all panel members to respond in a reasonably timely fashion for at least three rounds. Ideally, at least three months

should be allowed for three rounds to run their course. If the survey is relatively simple in design and e-mail can be used, then that three-month guideline can be dramatically contracted.

13. *Survey length.* As survey designers recognize, the number of people who will respond to a questionnaire is inversely related to the length of the questionnaire. However, in our research, it was impossible to avoid the length issue since a comprehensive inventory of applicable evaluation criteria for municipal plans was the desired end result. In all, more than 100 variables were listed in the preliminary rounds of the questionnaire. Nevertheless, by round 3, stable responses were achieved for nearly 30% of the survey variables, so the survey length actually contracted over time. Although no panel members specifically mentioned that the survey was too long, two panelists consistently commented on the time imposition involved. These observations were acted upon by minimizing the amount of supporting documentation that panelists were required to read, and by identifying some survey tasks as *strongly recommended* rather than *mandatory*.

14. *Establishing working relationships between the monitor and the panelists.* Individual attention to panel members contributes to satisfaction with the process. As well, the establishment of solid working relationships between the monitor and panel members can result in faster responses, and engender communications about future collaborations on subsequent research projects.

15. *Employment of a post-Delphi survey.* The post-Delphi survey is intended to provide insight into both the effectiveness and the delivery of the process, as well as to provide panel members with an opportunity to contribute their personal opinions regarding how the process was conducted. Panel satisfaction with results between rounds, the communication of information pertaining to the process, facilitation of learning, facilitation of participation, and overall performance of the Delphi monitor are all variables that can be addressed in the post-Delphi survey. The post-Delphi survey can accompany the final round, as was done in the case of our project.

Concluding remarks

Consensus-building techniques that employ information derived from public consultation can be used in many planning situations. However, if there is an unknown element to be explored, then the Delphi technique warrants consideration as the decision support instrument to use in the deliberation process. This paper contributes to the literature on Delphi methodology by discussing our experience in specifying and implementing a practical research design for a normative Delphi application.

Operational improvements to the Delphi process that were validated by our project include the following: the use of a pretest stage, the use of the precautionary principle as a stopping criterion for bimodality, the application of extremely strict panel selection criteria, and the employment of a graphic representation to visually demonstrate to panel members how the process is intended to progress. As well, lessons learned regarding the implementation of the technique are provided to promote and support the increased application of the Delphi in planning.

Planning as a profession is primed for the accelerated adoption of the Delphi technique, as the professional conduct and value of expertise components are already entrenched. Furthermore, urban planning issues are fraught with value-based complexities, and sometimes it is only the consideration of experts that can resolve the concern. While the Pythia of Delphi provided guidance induced by toxic fumes for war and agricultural scheduling, the scientific basis of the Delphi technique is grounded in methodological design. The continued viability of the technique in planning depends on a transparent substantiation of research design decisions, which was the focus of this paper.

References

- Ackoff R, 1953 *The Design of Social Research* (University of Chicago Press, Chicago, IL)
- Albright R, 2002, "What can past technology forecasts tell us about the future?" *Technological Forecasting and Social Change* **69** 443–464
- Cavalli-Sforza V, Ortolano L, 1984, "Delphi forecasts of land use: transportation interactions" *Journal of Transportation Engineering* **110** 324–339
- Chaffin W, Talley W, 1980, "Individual stability in Delphi studies" *Technological Forecasting and Social Change* **16** 67–73
- Coates J, 1999, "Boom time in forecasting" *Technological Forecasting and Social Change* **62** 37–40
- Coates V, Faroque M, Klavans R, Lapid K, Linstone H, Pistorius C, Porter A, 2001, "On the future of technological forecasting" *Technological Forecasting and Social Change* **67** 1–17
- Critcher C, Gladstone B, 1998, "Utilizing the Delphi technique in policy discussion: a case study of a privatized utility in Britain" *Public Administration* **76** 431–449
- Dajani J, Sincoff M, Talley W, 1979, "Stability and agreement criteria for the termination of Delphi studies" *Technological Forecasting and Social Change* **13** 83–90
- Dalkey N, 1969 *The Delphi Method: An Experimental Study of Group Opinion* (RM-5888-PR) (Rand Corporation, New York)
- De Boer J, Hale J, 2000, "The geological origins of the Oracle at Delphi, Greece", in *The Archaeology of Geological Catastrophes* Eds W J McGuire, D R Griffiths, P L Hancock, I Stewart (Geological Society, London) pp 399–412
- De Loë R, Wojtanowski D, 2001, "Associated benefits and costs of the Canadian Flood Damage Reduction Program" *Applied Geography* **21** 1–21
- Helmer O, 1983 *Looking Forward: A Guide to Futures Research* (Sage, Thousand Oaks, CA)
- Ilbery B, Maye D, Kneafsey M, Jenkins T, Walkley C, 2004, "Forecasting food supply chain development in lagging rural regions: evidence from the UK" *Journal of Rural Studies* **20** 331–344
- Jones C, 1975, "A Delphi evaluation of agreement between organizations", in *The Delphi Method: Techniques and Applications* Eds H Linstone, M Turoff (Addison-Wesley, Don Mills, ON) pp 160–167
- Linstone H, 1975, "Eight basic pitfalls: a checklist", in *The Delphi Method: Techniques and Applications* Eds H Linstone, M Turoff (Addison-Wesley, Don Mills, ON) pp 573–586
- Linstone H, Turoff M, 1975a, "Introduction", in *The Delphi Method: Techniques and Applications* Eds H Linstone, M Turoff (Addison-Wesley, Don Mills, ON) pp 3–12
- Linstone H, Turoff M, 1975b, "Evaluation: introduction", in *The Delphi Method: Techniques and Applications* Eds H Linstone, M Turoff (Addison-Wesley, Don Mills, ON) pp 229–235
- Linstone H, Turoff M, 2002, "The Delphi method: techniques and applications", <http://www.is.njit.edu/pubs/delphibook/index.html>
- Martino J, 1999, "Thirty years of change and stability" *Technological Forecasting and Social Change* **62** 13–18
- Masser I, Foley P, 1987, "Delphi revisited: expert opinion in urban analysis" *Urban Studies* **24** 217–225
- Nelson B, 1978, "Statistical manipulation of Delphi statements: its success and effects on convergence and stability" *Technological Forecasting and Social Change* **12** 41–60
- Novakowski N, 1999 *Derivation of an Evaluation Instrument for Judging the Quality of Ecosystem-based Municipal Plans* unpublished PhD dissertation, Department of Geography, University of Ottawa
- Parenté R, Hiöb T, Silver R, Jenkins C, Poe M, Mullins R, 2005, "The Delphi method, impeachment and terrorism: accuracies of short-range forecasts for volatile world events" *Technological Forecasting and Social Change* **72** 401–411
- Regier W, 1986, "Directions in Delphi developments: dissertations and their quality" *Technological Forecasting and Social Change* **29** 195–204
- Richey J, Mar B, Horner R, 1985, "The Delphi technique in environmental assessment: I. Implementation and effectiveness" *Journal of Environmental Management* **21** 135–146
- Rowe G, Wright G, 1999, "The Delphi technique as a forecasting tool: issues and analysis" *International Journal of Forecasting* **15** 353–375
- Scheibe M, Skutsch M, Scofer J, 1975, "Experiments in Delphi methodology", in *The Delphi Method: Techniques and Applications* Eds H Linstone, M Turoff (Addison-Wesley, Don Mills, ON) pp 262–287
- Sharma H, Gupta A, 1993, "Present and future status of system waste: a national-level Delphi in India" *Technological Forecasting and Social Change* **44** 199–218

-
- Turoff M, 1975, "The policy Delphi", in *The Delphi Method: Techniques and Applications* Eds H Linstone, M Turoff (Addison-Wesley, Don Mills, ON) <http://is.njit.edu/turoff>
- Turoff M, Hiltz S, 1996, "Computer-based Delphi processes", in *Gazing into the Oracle: The Delphi Method and its Application to Social Policy and Public Health* Eds M Adler, E Ziglio (Jessica Kingsley, London) pp 56–85
- Turoff, M, Li Z, Wang Y, Cho H, 2004, "Online collaborative learning enhancement through the Delphi method" *Proceedings of the OZCHI 2004 Conference, 22–24 November* University of Wollongong, Australia; <http://web.njit.edu/~turoff/Papers/ozchi2004.htm>
- Wellar, B, 1997, "Walking security index variables: initial specification", Interim Report 5, University of Ottawa and Region of Ottawa–Carleton, Ottawa
- Wellar B, 2005 *Geography and the Media: Strengthening the Relationship* (Canadian Royal Geographical Society and Canadian Council on Geographic Education, Ottawa)
- Wellar B, Vandermeulen C, 2000, "Field tests of the Driver Behaviour Index (DBI) survey forms: initial findings from an applied project involving selected regional intersections in Ottawa–Carleton", in *Papers and Proceedings of the Applied Geography Conferences* **23** 206–214
- Widstrand C, Kruus P, 1996 *Methods and Uses of Forecasting* 2nd draft edition (Carleton University, Ottawa)
- Wilhelm W, 2001, "Alchemy of the Oracle: the Delphi technique" *The Delta Pi Epsilon Journal* **43** 6–26
- Ying L, Kung J, 2000, "Forecasting up to year 2000 on Shanghai's environmental quality" *Environmental Monitoring and Assessment* **63** 297–312
- Ziglio E, 1996, "The Delphi method and its contribution to decision-making", in *Gazing into the Oracle: The Delphi Method and its Application to Social Policy and Public Health* Eds M Adler, E Ziglio (Jessica Kingsley, London) pp 3–33

Conditions of use. This article may be downloaded from the E&P website for personal research by members of subscribing organisations. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.