

# A Novel Approach to Reducing Uncertainty The Group Delphi

THOMAS WEBLER, DEBRA LEVINE, HORST RAKEL, and ORTWIN RENN

## ABSTRACT

A variation on the conventional Delphi was used to assemble an informational summary of expert opinion regarding the risks involved with the application of sewage sludge to farmland. The aim was to reduce uncertainties surrounding the associated health and environmental risks so agreement among citizens, farmers, and regulators could be reached. An expert panel was assembled for one day to take part in a structured communication process modeled after the Delphi. A two-part questionnaire using Likert scaling and open questions was iterated among rotating subgroups to build consensus and define disagreement. Plenary discussions were held between iterations to foster peer review.

There was consensus about the risks of heavy metals, pathogens, and nutrients; but clear disagreement about the risks of organic toxins. Existing state regulations were deemed inadequate only for lead and some organic toxins. Expert quantitative ratings were found to differ radically for two hypothetical contexts: academic and public.

## Statement of the Problem: Decision Under Uncertainty

Rarely do decision makers feel they have enough information available to make a decision about which they can be confident. Among regulators, politicians, managers, and public officials, considerable effort is expended in reducing uncertainty surrounding decisions. Uncertainty reflects incomplete knowledge. Depending on the type of knowledge, different techniques have proved effective at resolving uncertainties. Uncertain knowledge about factual evidence (random error, statistical error) can be resolved by scientific methodology and the peer review process, whereas uncertain knowledge about values (social priorities or preferences) can be reduced by political discourse or democratic voting procedures.

But the domains of these two procedures to reduce uncertainty are not inclusive to all kinds of knowledge. Uncertainties about predicting future events and uncertainties

---

THOMAS WEBLER is working on a doctoral dissertation at Clark University, Worcester, Massachusetts. DEBRA LEVINE is an analyst with the New York City Environmental Department. HORST RAKEL is a Fullbright Scholar from West Germany. ORTWIN RENN is an associate professor of Environment, Technology, and Society and a senior investigator at the Center for Environment, Technology, and Development (CENTED) at Clark University.

Address reprint requests to Prof. Ortwin Renn, ETS Program, Clark University, 950 Main Street, Worcester, MA 01610-1477.

The research reported herein was supported by a grant from the New Jersey Department of Environmental Protection. The results described in this paper do not necessarily reflect the viewpoints of the funding agency.

An earlier version of this paper was presented in August 1989 at the American Sociological Association conference in San Francisco.

about interpretations of factual evidence (theoretical concepts) cannot be resolved easily with either one of the above methods.

The uncertainty about predictions is close to the domain of science; it stems from the need to anticipate future conditions, that is, the parameters under which the scientifically revealed causal relations may operate. The problem of uncertainty then becomes one of how best to guess what the future conditions will be. Toward this end, the conventional Delphi has proved to be helpful. It uses a form of democratic group process to encourage experts to make a consensual "best guess" about future conditions.

The uncertainty about interpretations is close to the domain of political process. Much like the uncertainty about social values or preferences (resolved by a discourse), uncertainty about interpretations lies mainly in applying intrinsic values to evaluate factual evidence. For example, the question, What is pollution? requires a value choice about how clean the environment ought to be, but it also requires empirical evidence about toxicological thresholds. In this manner, interpretations link empirical concepts with theoretical concepts. To resolve uncertainty about this type of knowledge, the Delphi technique is practical, but only to a degree. If there is too much value conflict the process may be unacceptably elitist.

This paper examines a variation on the conventional Delphi as a method to reduce the uncertainty about environmental risks. The problem is to define the presumptions dominant within a specific expertise for the purpose of designing regulations to mitigate these risks. Identifying the risks involves resolving uncertainties about predictions (that is, What will the future conditions be?). Agreeing about the effectiveness of regulations at mitigating the risk involves resolving uncertainties about interpretations (that is, How safe is safe enough?). Among regulators, there is disagreement about factual evidence, disagreement about perceptions of probability (base-rate fallacy, subjective probability), disagreement about the effectiveness of various risk mitigation options, and disagreement about priorities. Often regulators are not aware of the current consensus (or lack of it) among the scientific community, or the regulations have not kept pace with the changing consensus.

In the case study, a variation of the conventional Delphi, called the Group Delphi, was used to structure a communication process in which a panel of experts discussed and defended their professional judgments concerning existing knowledge. The subject matter was the risks associated with the application of sewage sludge to farmland. The objective of the process was to evaluate the existing risk management regulations by defining when consensus was clear and characterizing the distribution of interpretations when consensus was not clear.

### **History of the Delphi Process**

Delphi is the name given to a procedure developed by the RAND Corporation in the 1950s to obtain consensus among a group of experts [1]. Like committee meetings, Delphi exploits the collective experience of the group members through an interactive process. The Delphi, however, avoids many of the pitfalls of conferences by structuring the communicative process in a unique manner. During the 1960s, Delphi became a popular tool for technological forecasting [2], and its applications have since expanded (see Sviden [3], for example). The definition used by Linstone and Turoff is still dominant:

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. [2, p. 3]

Often compared with committees and conferences, Delphi was distinct in the way

it (a) structured communication, (b) preserved anonymity of the participants, (c) gave feedback to the participants after each stage of the iterative process, and (d) produced a statistical group response [4, 5].

The conventional "Delphi exercise" is the basis for the many variations that have since evolved. The basic process involved a questionnaire that is designed by a facilitator or monitor team and mailed to a preselected participant group (expert panel). Upon its return, the monitor team summarizes the responses, develops a new questionnaire, and mails the summary and revised questionnaire to the same participants for another iteration. The experts are then able to reassess their opinions in light of other (anonymous) participants' comments and opinions, and make a revised judgment. The process continues until a steady state is reached. By refining the questionnaire, iteratively feeding information back into the process, and allowing the experts to change their positions, the extent to which consensus prevails is made clear.

### **Shortcomings of Delphi**

#### SELECTION

The problem of selecting experts is twofold: identifying the experts and selecting who participates. While errors in the questionnaire can be corrected, an incapable or misrepresentative expert panel will compromise the process. To achieve legitimate results it is important that all relative viewpoints are represented [4].

The central concern in assembling the expert panel is to represent the entire array of perspectives within the discipline. It is not necessary to have a representative sample, but the selection should be systematic. Differing schools of thought can be identified through journals or through conversations with several people knowledgeable in the discipline. While it is useful if at least one member of the monitor team has a working knowledge of the specific discipline (experts who feel the Delphi organizers are unfamiliar with the subject matter will not participate), too much expertise may imply bias.

Another aspect to consider in the selection of the Delphi panel is cultural bias. Questionnaires unavoidably rely heavily on subjective definitions. If experts do not share the same biases, the questions will be interpreted differently. The only way to mitigate this distortion is to ensure a diversity among the panel members so that the biases cancel out [6].

Systematic biases among the expert panel should not be ignored. Overconfidence, belief in quantification, and insensitivity to sample size are three examples of expert bias that, if present, will be reflected in the results [7, 8].

#### PARTICIPATION

Experts are busy people. A questionnaire that requires substantial effort may not reach the desk of someone not directly involved in the matter. On the other hand, experts who have a stake in the particular subject are more willing to participate, but their opinions may be overly biased.

The dropout rate of participants is typically high [9]. This may be due to their underestimating the work involved, or because the task was inaccurately described by the Delphi organizer. In a circumstance where a particular expert viewpoint is weakly represented on the panel, a substantial dropout rate could sacrifice the validity of the process.

#### INTERPRETATION OF RESULTS

Delphis, by design, ask questions that require a large degree of subjective opinion. This makes them vulnerable to heuristic biases. Forecasts are subject to several distortions:

discounting the future, the prediction urge, and the simplification urge [ 2, pp. 574–580]. Opinions about the relevance of new information may be distorted by biases in the representativeness heuristic [7].

A key weakness of any Delphi procedure is the possibility that the organizer may be deceptive. Cyphert and Gant showed how purposely inserted fallacious information can affect the responses of Delphi participants, altering the voting distribution [10]. Inasmuch as the moderator receives and edits all comments and responses, and feeds these back into the process under a cloak of anonymity, the moderator is ideally situated to deceive. This shortcoming, however, is in no way unique to Delphis. Perhaps more likely is the possibility that the moderator may mistakenly misinterpret the responses while summarizing the results.

#### APPLICABILITY AND COMPETENCY

Delphi has been used extensively to resolve uncertainty about future conditions [2]. Not simply guesswork, the expert panels were asked to use known cause–effect relations to extrapolate conditions of likely scenarios within scopes of predefined epistemological frameworks, such as economic, behavioral, political, and technical. Benarie recently reported that many processes of environmental standard setting are “Delphi-like” [11].

As Martino has pointed out, the question of accuracy of Delphi results is inappropriate [12]. Rather, the correct question is, Does the process capture the expert’s views? A Delphi is a petition for expert *opinion* about anticipated future conditions or interpretations of factual evidence. Accuracy is characteristic of factual evidence, uncertainties about which are resolved with scientific method.

The Delphi process is not suited to all types of problems. It is best suited for those that involve a mixture of scientific evidence and social values. Delphi offers a flexibility along with a validity that is difficult to achieve with conventional strategies in problems of this mix. Linstone claims Delphi is particularly useful when holding a conference is impractical (because of cost, time, or domination of strong personalities) or when

The problem does not lend itself to precise analytical techniques but can benefit from subjective judgments on a collective basis. [5, p. 627]

Besides technological event and trend forecasting, the Delphi process has been used to evaluate budgets, define policy options, expose hidden agendas, and assess the significance of past events. Linstone clearly indicates that he expects Delphi’s applications to grow:

Perhaps the most important advice to be given to the potential user is to suit the method to the problem, not the problem to the method. [5, p. 627]

#### **A Modified Delphi Exercise: The Group Delphi**

The Group Delphi, or Expert Workshop, is a variation on the conventional Delphi exercise designed to consolidate expert opinion within a particular discipline in a very short time period. The Delphi technique of repetitious questionnaires and feeding results back into the process is exploited to encourage consensus about issues; however, the aspect of anonymity is given up. Scaling is employed to define deviant opinions. In the Group Delphi the feedback process is conducted as a conference. During a plenary session the moderator asks the experts to justify or make a case for their opinion, and discussion among the panel is encouraged [13, 14].

#### SELECTION OF THE EXPERT PANEL

As in other Delphi processes, considerable time is spent selecting the expert panel. Restraints are more severe for the Group Delphi because of the need to physically gather

the panel in one location at one time. Because an important aspect of the Group Delphi is the plenary sessions during which experts with deviant opinions are asked to defend their stances, it is thought that video or telephone conferencing cannot serve as an acceptable substitute for physical attendance. To ensure efficient communication during the plenary, the number of participants must be limited based on the amount of time available. A one-day Group Delphi may encompass three or four iterations. For such an encounter, an expert panel of 10–20 members is optimal.

#### THE PROCESS

The Group Delphi process is extremely efficient at bringing about consensus because of its two-tiered structure. Experts are initially brought together in a plenary where the process is introduced and questions are answered. Next they are divided into several small groups (three to five groups, each with three to four members) and given the questionnaire. Each small group works in a private room and is instructed to try and reach consensus on each question, although majority/minority votes are allowed in the first round. Upon completing the questionnaire, a break is usually scheduled to give the moderator time to assemble the results. Once this is completed, the plenary is reassembled and the moderator presides, systematically reviewing the questionnaire results, identifying deviations, and asking the subgroups to justify their positions. The moderator permits discussion when such discussion appears helpful in having the group reach consensus, but when it is clear that two camps are firmly established, the moderator redirects the group focus to the next item.

In the second round, the membership of the small groups is shuffled. There may have been a few changes made to the questionnaire during the plenary, but there is no time to redesign the entire questionnaire. The results of the second round are assessed as before and the plenary review process is repeated. This tends to move much faster as, on many points, consensus is achieved in the first round.

Linstone hinted of such an application when he viewed the Delphi process as a communication process and noted:

The efficiency of face-to-face meetings can be increased by a supplemental group communication process. [2, p. 4]

#### COMPARISON TO CONVENTIONAL DELPHI

The major elements of the conventional Delphi are iteration with feedback of responses into the successive round, assessment of group judgment, anonymity of participants, and opportunity for participants to revise their views. Group Delphi preserves all but anonymity.

Anonymity is considered important to ensure that the experts give their own professional opinions and are not swayed by the personalities of the other panelists. A criticism of conferences is that forceful personalities may intimidate the shy or uncertain participants into joining their side. This “bandwagon effect” can produce an illicit consensus. One approach to solving this problem with hardware is William Simmons’s Consensor—an electronic device that is designed to reduce this problem by enabling meeting participants to express their agreement or disagreement anonymously [15]. While this has been found to be useful at improving the efficiency of meetings, it does not eliminate the need to surrender anonymity when there is disagreement among the responses. Although giving up anonymity is not essential to the Group Delphi, a satisfactory and practical technique to preserve it during the process has not been found. In short, anonymity is surrendered to enhance the consensus-building procedure.

To minimize the likelihood of the bandwagon effect occurring, experts of equal

status should be chosen and, during the selection process, particularly domineering persons should be weeded out. Even after the event has begun, some of the problematic personality differences can be negated by subtleties of the moderation. The moderator can directly ask an untalkative panelist to contribute an opinion, or quiet those who are too outspoken. There is still the possibility, however, that the artificiality of the situation may cause the less confident experts to say what they think others want to hear, and not to give their honest opinion.

There are three major advantages to having face-to-face communication. First, a chief criticism of the conventional Delphi is that information from respondents may be distorted, intentionally or unintentionally, by the moderator; because views are discussed openly in the Group Delphi, there is direct and immediate feedback. Any ambiguities are immediately clarified. Second, the justifications given for dissenting viewpoints also give secondary insights into which deviations are accepted by the panel. Third, these discussions provide an internal check for consistency in accepted viewpoints.

The last major difference between the Group Delphi and the conventional Delphi is time. The Group Delphi is accomplished in one or two days, whereas the conventional Delphi may take several months. If we view the process as a measurement, then the Group Delphi gives a single data point while the conventional Delphi provides a running average. The Group Delphi, then, can record more spurious reactions. If there occurred a substantial event related to the subject matter, then a series of Group Delphis around the event could reveal how it affected expert opinion. This may be a disadvantage if the purpose of the Group Delphi is to make a baseline measure and the process is swayed by an unforeseen stimulus.

Another critique is that this short operating time restricts the feedback to simple statements. Information must be quantified to do this assessment quickly. Furthermore, there is no time to redesign the questionnaire between rounds, although the panel could be convened on another day.

### **Case Study: Land Application of Sewage Sludge**

#### **INTRODUCTION**

A Group Delphi was organized to reduce uncertainties about the risks involved with applying sewage sludge to farmland. The purpose was to provide regulators (or, as in this case study, a panel of public citizens) with a summary of expert opinion in the discipline. Uncertainties about causal relationships (predicting leaching, runoff, crop uptake, and so on, of hazardous compounds) and uncertainties about interpretations (theoretical concepts, that is, How safe is safe enough?) were addressed using the Group Delphi process. The goals were to (a) identify the hazards, (b) rate the relative risks associated with the hazards, (c) judge the effectiveness of existing regulations at mitigating the risks, and (d) suggest improvements to the regulations to achieve acceptable protection.

#### **METHODS**

##### *Selection of Participants*

The selection process began with the names of experts chosen from recent books and articles published on the subject. To establish expertise, a self-rating scheme was used, whereby the experts were asked whether they considered themselves to be experts. They were also asked to recommend other experts whose presence would complement the Delphi panel.

There was a clear consensus among those familiar with current research in the subject that there were two rival camps, each with their own leader. One of the leaders was

adamant about having his opponent attend. He apparently wanted to use the opportunity to debate. While it may be disadvantageous to have two strong personalities monopolize the conversation, a well-structured process should be able to keep such debate under control. Unfortunately, the other leader did not want to debate; he resisted all reasonable invitations to participate. We were successful, however, at having representatives of both camps attend and they proved capable of arguing for their opinions.

In sum, 34 persons were contacted, and 21 were interested and considered themselves capable. Five dates were presented for holding the workshop, and those who were interested were asked which dates they could not attend. From this, the optimal date was chosen, 13 experts were invited, and 10 promised to attend. One expert fell sick the day before the event; in all, nine experts attended.

Participating were an EPA researcher; university researchers from various universities representing the following disciplines: chemistry, crops and soils sciences, horticulture, and land reclamation; city water and sewage authorities; a consultant; and a private contractor who applies sludge.

### *Design of Questionnaire*

The questionnaire was composed in two sections. The first section asked the experts to rate the risks to humans and the risks to the environment associated with specific compounds likely to be found in typical municipal sewage sludge that met the regulatory criteria for land application. To incorporate the experts' uncertainty into the evaluation, confidence ratings were petitioned for each answer. Except for limits on the maximum concentration of contaminants in the sludges, no other risk-reduction regulations were meant to be considered in this section. A quantitative scaling method from 0 to 9 was used for all questions.<sup>1</sup> The experts were told to calibrate the scale so that 9 referred to a risk  $> 10^{-3}$  and 0 to a risk of  $< 10^{-6}$ . Under each category were blank spaces which experts could use to write in other risks they deemed important.

Questions in the second section asked the respondent to rate the effectiveness of specific regulations at mitigating the risk.<sup>2</sup> Again, each question had two thrusts, one toward risks to humans, another for risks to the environment. With each rating, the experts were asked to give an estimate of how confident they felt (in percent) about their numerical responses.

### RESULTS

The Group Delphi process proved efficient at bringing about consensus. After two rounds it was clear that there was agreement on most items. Where disagreement occurred, both camps argued for their positions.

One completely unexpected result was a clear distinction between quantitative ratings

<sup>1</sup> Risk assessments										
Compound	Environmental risks			% Confidence						
				Health risks	% Confidence					
PCBs	0	1	2	3	4	5	6	7	8	9
PAHs	0	1	2	3	4	5	6	7	8	9
Dioxins	0	1	2	3	4	5	6	7	8	9

<sup>2</sup> Suitability of regulatory requirements																	
Standard	Too lenient			Too strict	Effectiveness for risk reduction	Proposed standard											
Cd 20 ppm	-3	-2	-1	0	+1	+2	+3	0	1	2	3	4	5	6	7	8	9
Cu 600 ppm	-3	-2	-1	0	+1	+2	+3	0	1	2	3	4	5	6	7	8	9

that experts gave for risks under two hypothetical audience contexts: academic and public. Risks that were rated slight, such as 2 or 3 (on a scale of 0 to 9, 9 being great risk), in an academic setting would be rated a 0 (no risk) in a public setting. Likewise, where the academic context accepted confidence levels of < 100%, experts felt that the public would accept no statements that were not accompanied by 100% confidence.

Responses from the first round showed extreme disagreement to the point that we suspected two participants of willingly trying to corrupt the outcome. During the plenary, they were asked to defend their responses and it was learned that they were trying to make a point. Their point was that animal manures posed a much greater risk because they are not regulated. Most of the panel agreed with this point, but had answered the questionnaire under the assumption that it referred to the relative risks each compound posed. Those who gave low risk ratings to sewage sludges had calibrated the scale according to other comparative risk sources such as manure. Since the entire panel had difficulty distinguishing between risks to humans and risks to the environment, they claimed the distinction was unnecessary, we adjusted the questionnaire for the second round by replacing the column heading "human risks" with "relative risks" and replacing "environment risks" and "absolute risks." The convergence of response upon the second round showed this correction to be effective at resolving the confusion.

The panel agreed about the severity of risk associated with heavy metals, nutrients, and pathogens. There has been substantial research into these compounds, and they are thought to be well understood. Regulations designed to mitigate the risks associated with heavy metals, nutrients, and pathogens were considered sufficient or too strict—with one exception. Lead was not regulated strictly enough. The regulations allowed lead in concentrations up to 4800 parts per million (ppm). The levels suggested by the three subgroups of experts during the second round were very similar: 800, 1000, and 1200 ppm.

On the topic of organic toxins, there was clear disagreement. In fact, this is the subject in which the two leading sludge researchers (mentioned above) disagreed most strongly. One camp thought the risks of organic toxins to be extremely high and faulted the regulations for being incomplete and too lenient. For example, the regulations permit polychlorinated biphenyls (PCBs) at concentrations of 0.5 ppm in the sludge. One expert claimed that 1 part per billion (ppb), if consumed, is cancer promoting, and some root crops will take up PCBs in concentrations nearly equal to that of the soil. He wanted the regulations to be more strict. Others countered that PCBs are ubiquitous and the amount contributed by sewage sludge to the total soil concentration is insignificant. Another argument was that polyaromatic hydrocarbons (PAHs), known to be in sludge, have recently been determined to be carcinogenic; however, they are not regulated at all. Because recent data appeared to indicate that these compounds are more dangerous than originally thought, some experts suggested that land application regulations require extensive testing for these organic toxins.

The other camp admitted that toxic organic compounds, including PAHs, are poorly understood, largely because of their sheer numbers (hundreds of thousands) and the high cost of testing for them; but they found the risks acceptable because the level of toxicity was thought to be on the same order of magnitude as the surrounding soil and air. There was a consensus, however, that the main focus of future research should be to study the toxic organic compounds in sludges and the risks associated with applying these sludges to food crops.

Figure 1 depicts the development of consensus from round one to round two. For metals, pathogens, and nutrients, the confidence levels for the absolute risk ratings improved and the variance among the three groups decreased. Where there had been agree-

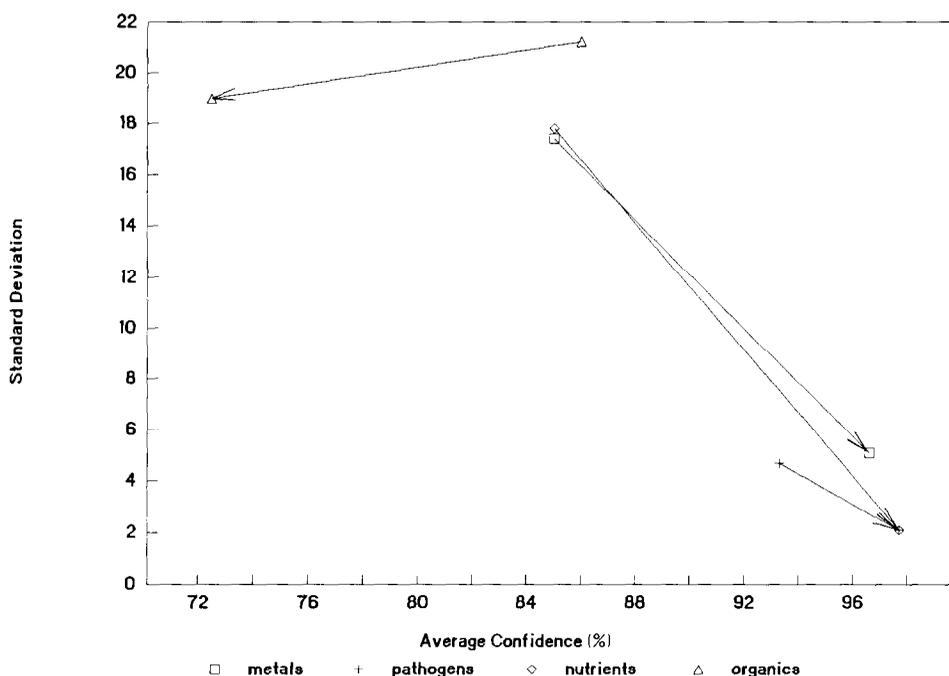


Fig. 1. Change in average confidence and standard deviation for absolute risk ratings from round one to round two. Responses of all three expert groups are combined for each item in the four contaminant types.

ment on the risks posed by these compounds in the first round, the process improved this agreement. Organic toxins, on the other hand, remained a topic of contention. As a consequence of the discussion that followed the first round, the experts reevaluated their confidence ratings, indicating there was more uncertainty about the risks of these compounds than they originally suspected. These results show that the process was effective at clarifying disagreement as well as agreement, and that the lack of anonymity did not bring about a coerced consensus.

It was clear that each expert on the panel had a specific expertise about which he or she felt comfortable speaking. Most kept to their specialties, addressing only two or three issues. The organic chemist did not speak at all until the issue of organic toxins arose, then he spoke at great length. The moderator actively elicited the opinions of the more quiet panelists. By encouraging others to participate and even calling for specific expert perspectives, the moderator was able to produce a balanced dialogue.

#### DISCUSSION

Judging by the plentiful debate in the plenary sessions and the satisfaction of the participants as the event ended, the selection process did assemble a capable expert panel. There was mutual respect among the experts, which indicated they were of similar status. The nonattendance of the opposing "kings" did not cripple the discussion.

The outcome of the event had expected and unexpected aspects. As expected, uncertainty about risks was reduced. On a large majority of items consensus was found, but there was also clear disagreement on some issues. Importantly, we were able to sort out which was which. Regarding compounds that have been well researched and have a

long history of familiarity (pathogens, metals, nutrients), the experts indicated there was little uncertainty about future conditions. Interpreting the level of acceptable risk for these compounds also did not split the panel. Regarding the organic toxins, which are much less specifically understood (because of newness and multitude), the expert panel split on opinions of future risks and levels of acceptable risk. Since questions of acceptable risk require interpretations of evidence under specific value preferences, the inability of the panel, as a whole, to come to consensus on this issue suggests that the value preferences of society have not been clearly defined.

The unexpected outcome was that the perspective taken by an expert (the role one is asked to play) affects the ratings assigned to various risks. When addressing the public, experts feel a need to be definitive, "low risk" becomes "no risk," and confidence ratings increase to 100%. This suggests two things: first, that the experts feel the public does not completely understand scientific rationality. Whereas the academic community would scorn a member who claimed 100% confidence about a value-laden issue, the experts in our study felt that this honesty with the public would delegitimize the scientific contribution. Second, it suggests that the experts feel it is extremely important that their evidence and opinions be incorporated into the public debate, or else they would not put such emphasis on anticipating public reaction. The absence of anonymity did present a problem: the plenary discussions were dominated by a few people. The problem was manageable, however, and did not prohibit discussion of confrontational issues.

The questionnaire had two characteristics: a standardized format and open questions to encourage unforeseen responses. Standardization of the questionnaire is necessary to define deviations and to structure the discussion. Open questions are extremely important because there is no time to rewrite the questionnaire from one round to the next. One concern noted by the experts in the open question was that property values may decline for homes adjoining farms using sewage sludge. This was not foreseen in the original questionnaire.

This event took place on one day. It provides only a snapshot of the expert opinions on that day. As new research is completed and more experience gained, these opinions are likely to change. Therefore, it may be necessary to arrange another Group Delphi in the following year, to keep abreast of the changes.

The purpose of this Delphi was to provide information for a public participation project. The results produced were suitable toward that purpose. There were two products: the completed questionnaire and videotapes. The completed questionnaire was input to the public participation project, and the videotapes were made during the final plenary to record the discussion for secondary analysis.

## **Conclusions**

The Group Delphi is an effective technique for reducing uncertainty surrounding knowledge about predictions and interpretations. It brings a convenience in both time and effort over the conventional Delphi, but, in doing so, provides only a snapshot of the expert opinion in the subject. Like the conventional Delphi, a heavy emphasis is placed on assembling an expert panel that represents all points of view. Unlike the conventional Delphi, it demands that the anonymity of the panel be given up, in exchange for face-to-face communication. By using a two-tiered structure of small and large groups, the Group Delphi is extremely efficient at defining areas of consensus, but is limited in size and scope. Because of time restrictions, the questionnaire must be flexible and standardized if the process is to function effectively.

There are three advantages of the Group Delphi not associated with the conventional Delphi:

1. It produces a clearer picture of the dissention among the expert panel.
2. There is justification given for the dissent.
3. There is direct testing of dissent in a "peer review."

In the case study, the Group Delphi was effective at defining expert opinion about risks and regulations. We discovered that the experts rated risks differently in a closed academic session than they would in a public setting.

As a technique to reduce uncertainty, the Group Delphi has shown promise. It is an example of how the conventional Delphi can be altered to include argumentation for different viewpoints.

*The authors thank Thomas Dietz for his helpful comments on an earlier version of this paper.*

## References

1. Dalkey, N., and Helmer, O., An Experimental Application of the Delphi Method to the Use of Experts, *Management Science* 9, 458–467 (1963).
2. Linstone, H. A., and Turoff, M., eds., *The Delphi Method*, Addison-Wesley, Reading, MA, 1975.
3. Sviden, O., Future Information Systems for Road Transport: A Delphi-Derived Scenario, *Technological Forecasting and Social Change* 33, 159–178 (1988).
4. Martino, J. P., *Technological Forecasting for Decisionmaking*, Elsevier, New York, 1972.
5. Linstone, H., The Delphi Technique, in *Handbook for Futures Research*, J. Fowles, ed., Greenwood Press, Westport, CT 1985.
6. Hill, K. Q., and Fowles, J., The Methodological Worth of the Delphi Forecasting Technique, *Technological Forecasting and Social Change* 7, 179–192 (1975).
7. Tversky, A., and Kahneman, D., Judgment Under Uncertainty: Heuristics and Biases, *Science* 27 September 1974, p. 47.
8. Slovic, P., Fischhoff, B., and Lichtenstein, S., Rating the Risk: The Structure of Expert and Lay Perceptions, in *Environmental Impact Assessment, Technology Assessment, and Risk Analysis*, V. T. Covello et al., eds., Springer-Verlag, Berlin, 1985.
9. Jillson, I. A., The National Drug-Abuse Policy Delphi: Progress Report and Findings to Date, in *The Delphi Method: Techniques and Applications*, H. A. Linstone and M. Turoff, eds., Addison-Wesley, Reading, MA, 1975.
10. Cyphert, F., and Gant, W., The Delphi Technique, *Journal of Teacher Education* 21, 417–429 (1970).
11. Benarie, M., Delphi and Delphilike Approaches with Special Regard to Environmental Standard Setting, *Technological Forecasting and Social Change* 33, 149–158 (1988).
12. Martino, J. P., *Technological Forecasting for Decision Making*, North Holland, New York, 1983.
13. Renn, O., and Kotte, U., Umfassende Bewertung der vier Pfade der Enquete-Kommission auf der Basis eines Indikator Katalogs, in *Energie im Brennpunkt*, G. Albrecht and H. U. Stegelmann, eds., HTV, Munich, 1984.
14. Renn, O., Decision Analytic Tools for Resolving Uncertainty in the Energy Debate, *Nuclear Engineering and Design* 93, 167–179 (1986).
15. Simmons, W., The Consensor: A New Tool for Decision-Makers, *The Futurist* 13, 91–94 (1979).

*Received 30 October 1989; revised 15 May 1990*