Technical Report No. 22

EVALUATION OF METHODOLOGY IN THE UNIVERSITY OF MICHIGAN'S SEA GRANT DELPHI INQUIRY

By
JOHN D. LUDLOW
Assistant to the Director
University of Michigan Sea Grant Program

February, 1972
MICHU–SG–72–202

Multidisciplinary Research in the Great Lakes

EVALUATION OF METHODOLOGY

in the

UNIVERSITY OF MICHIGAN'S

SEA GRANT DELPHI INQUIRY

John D. Ludlow
Assistant to the Director
University of Michigan Sea Grant Program

February 1972

MICHU SG TR# 22
Sea Grant Technical Report No. 22

MICHU-SG-72-202

THE UNIVERSITY OF MICHIGAN SEA GRANT PROGRAM

The University of Michigan Sea Grant Program is a part of the National Sea Grant Program, which is maintained by the National Oceanic and Atmospheric Administration of the U. S. Department of Commerce.

ABSTRACT

This paper evaluates the effectiveness of the Delphi techniques in solving two related management problems: (1) integrating the judgments of an interdisciplinary research team, and (2) conveying its informed insights to regional decision makers. The substantive results of the Delphi exercises are covered in detail in another Sea Grant report.

BACKGROUND

This study was funded by the Sea Grant Program of the University of Michigan and conducted by the Bureau of Business Research, Graduate School of Business Administration, the University of Michigan. We are greatly indebted to the sixty-nine University of Michigan researchers and concerned citizens from the Grand Traverse Bay area who contributed to the substantive results of the Delphi exercises and to an evaluation of the methodology.

CONTENTS

TABLES

FIGURES

I

INTRODUCTION

The development of methods to obtain and refine the informed judgments of knowledgeable people is one of the most crucial problems in planning and decision making. The task is particularly challenging in the Sea Grant Program, which emphasizes a systems approach by a multi-disciplinary group of researchers--some of them experts in extremely specialized areas, representing a wide range of technical, economic, social, legal, and political matters. The Sea Grant Program of the University of Michigan provided an unusual opportunity for a critical evaluation of the Delphi techniques in an operational environment.

The interdisciplinary breadth of the Sea Grant Program has been a central question to the Program Committee, that is, how to mesh disparate disciplines successfully so as to secure the full strength of their competence as integral parts of a program of long-range research and planning for water resource management. The Committee recognized that the development of integrative functions constituted a departure from existing academic activity.[1/]

1/ Proposal to National Oceanic and Atmospheric Administration (NOAA), "Continued Participation in the Sea Grant Institutional Support Program, 1971-72," Vol. II, submitted by the University of Michigan, Jan. 1971.

Once a model indicating the impact on water resources of alternative planning strategies is established, society must be informed through methods brought to light by the present program in order for choices to be made through some acceptable political process.[2/]

The Delphi exercises were designed to support the Sea Grant goals of encouraging (1) the involvement of university people in a concern for comprehensive management of the water resources of the Great Lakes, (2) the integration of their informed judgments, and (3) the communication of these judgments to citizens who are faced with decisions that will affect their region's social and economic development for many years.

The Basic Delphi Method

The Delphi[3/] method evolved from a series of studies conducted by researchers (initially under the guidance of Olaf Helmer and Norman Dalkey) of the RAND Corporation. The basic method was defined as a

---

2/ Ibid.

3/ The name suggests the practices of the ancient Greeks of obtaining the counsel of a deity through oracles, one of the most renowned of which was the oracle at Delphi.

set of techniques for soliciting and collating the opinion of experts in order to arrive at the most reliable consensus. The method's most distinctive features are anonymity, statistical summaries of information provided by the group, controlled feedback, and an iterative process that permits and encourages the reassessment of initial judgments.

Anonymity is observed in regard to specific inputs but not necessarily in regard to panel membership. Scaled descriptors are used to quantify judgments regarding such things as personal assessments of probability, importance, economic and technical feasibility, desirability, and the assessor's own competence. Statistical summaries ensure that all judgments are represented and provide a mechanism to identify extreme positions for which brief supporting assessments are usually requested. Advocates of the Delphi techniques generally claim that they preserve the desirable features of committee meetings while reducing some of the associated behavioral and administrative problems, and that they result in the most reliable measure of group opinion.

Conceptual Modifications

For the Sea Grant exercises expert opinion was transformed into informed judgments. This provided the rationale for the inclusion of

politicians and concerned citizens on the panels and exploited the inherent characteristic of the method: to inform during the process of soliciting judgments. The behavioral sciences support this type of approach to effective communication. For example, McGregor states:

> ...my conception of a two-way communication is that it is a process of mutual influence. If the communicator begins with the conviction that his position is right and must prevail the process is not transactional but coercive.4/

and similarly from Drucker:

> There will be no communication, in sum, if it is conceived as going from "I" to "Thou." Communication only works from one member of "us" to another.

> They must understand it because they have been through it, rather than accept it because it is being explained to them.

> In a complex world there is a need for a shared experience in the decisions, or there is no common perception, no communications, and, therefore, neither acceptance of the decisions,

---

4/ Douglas McGregor, The Professional Manager (New York: Harper & Row Publishers, 1967), p. 153.

> nor ability to carry them out. The ability to understand presupposes prior communication. It presupposes agreement on meaning....5/

The concept of consensus was shifted from an emphasis on unanimity and a meeting of the minds to a measure of collective judgments. The utility of a narrow or strong consensus on specific issues was recognized, since regional planning is basically a political process and some of the panelists are faced with decisions regarding many of the issues considered. However, narrowing the measure of agreement should be a result of reassessments based on such things as the consideration of additional information and related issues or events--not on some form of manipulation. Finding areas of uncertainty and underlying reasons for disagreement was important to the entire program, since the Delphi exercises were intended to provide information to other projects. They were designed as a complement to--rather than a substitute for --interpersonal techniques for securing and integrating subjective judgments.

5/ Peter F. Drucker, Technology, Management, and Society (New York: McGraw-Hill Book Co., 1970), pp. 22-23.

II

OVERALL PLAN

Three groups, formed on the basis of their training and experience, participated in the iterative cycles of feedback and reassessments that characterize the basic Delphi method. Two of the groups were composed of academicians and the third group was made up of concerned citizens from the Grand Traverse Bay area.

The first group was people whose expertise is primarily in the physical sciences and who were equally divided between Sea Grant researchers and experts from the School of Engineering. They are in technical areas that are relevant to water resource management. This group constituted the technical panel and was referred to as technicians.

Another panel was made up of the balance of the Sea Grant researchers who were not members of the technical panel. Generally, their academic backgrounds and interests are oriented toward the social sciences and the group was labeled behaviorists.

A third panel represented a group of persons from the Grand Traverse Bay area who are believed to be influential in the political processes through which decisions regarding regional planning are made. This group was called decision makers.

Although the names associated with the groups are somewhat arbitrary and the groups are far from homogeneous, the designations of technicians, behaviorists, and decision makers are reasonably consistent with the roles each would be expected to assume in the deliberations and actions associated with planning for future regional development.

The process of information feedback and reassessment of estimates was intended to be cumulative--each series building on the other and guiding the respondents toward carefully formulated judgments. Specialized issues and specific developments were considered in the early rounds; the stage was thus set for group recommendations regarding research priorities and planning policies in the final rounds.

The administrative procedures were flexible to allow modification and refinement of the methodology to be cumulative also. For example, suggestions of the technical panel were integrated into the exercises involving the broader-based panels. A progressive type of Delphi was planned so that important developments in the technical environment could be generated and assessed by the technical experts before being considered by broader multidisciplinary panels. Regional planners and decision makers were included in the broader-based panels, not only to capitalize on their knowledge of the region and the issues but also to provide a systematic and convincing way of giving them information that may be useful in regional planning.

Specific Objectives and Methodology

When the Sea Grant exercises were planned the Delphi method had been principally used for generating a list of technical events and soliciting subjective judgments relative to them. One phase of the Sea Grant Delphi exercises was devoted to identifying potential technical, social, and political developments. This phase will be described in enough detail to provide a review of the basic Delphi method. 1/ Other phases designed to accomplish specific objectives will be mentioned only briefly.

Potential Developments and Requisite Technology

Each of the three panels was provided with four rounds of information packages (instructions to the panel, published articles, evaluations matrixes, statistical summaries, etc.) which included a form for suggestions and estimates about future events. In the opening round respondents were asked to suggest potential developments and requisite technical events that they felt would be important to the marine

---

1/ For a more complete discussion of the basic method see The Delphi Method: A Systems Approach to the Utilization of Experts in Technological and Environmental Forecasting, by John D. Ludlow, Working Paper No. 3 (Ann Arbor, Mich.: Bureau of Business Research, Graduate School of Business Administration, University of Michigan).

resources of the Grand Traverse Bay area in the next twenty years. They were not provided with a list of developments by the administrator, a common way of eliciting items to be considered and facilitating participations. Instead they were asked to list their suggested developments and to evaluate several items according to factors on the evaluation matrix in order to familiarize themselves with the numerical estimates that would be requested on subsequent rounds.

Collating the developments suggested by panel members resulted in sixty-two discrete items. These were fed back to the panel on the second round along with thirteen events that were taken from other Delphi exercises conducted at RAND and the Institute for the Future. This latter group of items was included because the researcher felt that they dealt with areas which would be of interest to the panel and also because they were considered good examples of how developments should be specified to avoid ambiguity, particularly regarding occurrence or nonoccurrence.

The evaluation matrix for the second round requested self-evaluation estimates about familiarity with the Grand Traverse Bay area and with each item being considered. Respondents were also asked to make estimates of importance, economic and technical feasibility, and the probability and timing associated with the developments

with which they had some familiarity. The probability estimates dealt with the likelihood that events would occur during a specific planning period (1971-80) and also for time periods associated with specific probabilities (25 per cent, 50 per cent,and 75 per cent).

Panelists were asked to suggest groups of listed developments that were related and requisite technology associated with specific developments. Further clarification of several suggested requisite technical events was requested so that the events could be phrased appropriately for the type of numerical estimates which were being solicited in the Delphi exercise.

In their responses to the information package for the second round, some panel members indicated that making so many estimates took too much time and became tedious. The items to be included on the third round were therefore carefully screened and the evaluation matrix was considerably simplified.

Items were dropped from further consideration during the screening process if the average estimates of the panel indicated any of the following conditions:

a. The item was unimportant.
b. The panel was unfamiliar with the item.
c. The specification of the item was ambiguous.
d. A reasonably strong consensus existed regarding the item.

As the rounds progressed the feedback of information from the previous round became more comprehensive. Statistical summaries were [illegible]

which were based on the responses of the entire group of technicians, the subgroup who rated their familiarity with the item relatively high, and those in the subgroup who also rated their familiarity with the Grand Traverse Bay region relatively high. Arguments to support extreme positions were included.

On the final round the technicians were asked to make additional probability estimates for pairs of events that panel members had suggested were closely related. For these estimates they considered both the occurrence and nonoccurrence of the related (conditioning) event, that is, whether the occurrence or nonoccurrence of one event would affect the outcome of a following event.

The broader-based panels considered many of the items that had been evaluated by the technical panel and additional items that were suggested by broad-panel respondents as being of special interest to them. Several innovations designed to improve communication between researchers and decision makers were introduced and are discussed under methodological modifications (Chapter III).

Social, Political, and Economic Indicators

The panel was asked to project the trend curves of statistical values that are ordinarily used to measure the socioeconomic growth

of a region.[2/] This portion of the Delphi exercises was designed to provide panelists with points of reference in making other subjective judgments.

Relative Importance of Pollution Sources

Using an iterative procedure, the technical panel established a list of pollution sources for two time periods (1971-80 and 1981-90). These sources were then ranked according to their importance. The estimates for a body of water having the general characteristics of Grand Traverse Bay were conditioned by a systematic consideration of specific pollutants and the economic and technical feasibility of their abatement. The ten most important sources of pollution in the judgment of the technicians were submitted to the broader-based panels for their consideration.

Recommendations Regarding Waste Water Treatment and Disposal Systems

Recommendations relative to waste water treatment and disposal systems were solicited from the technical panel after they had considered

---

2/ The techniques and procedures used in this series of interrogations are similar to those described in Raoul de Brigard and Olaf Helmer, Some Potential Societal Developments--1970-2000. IFF Report R-7 (Middleton, Conn.: Institute for the Future, Apr. 1970).

the relative importance of pollution sources, as well as important social, political, and technical developments. The evaluations were designed as a risk-benefit-cost type of analysis. The procedures were similar for the broader-based panels.

Regional Planning Considerations

The judgments for regional opportunities and problems, regional planning strategies, and research and information priorities were solicited on the later rounds. This was done so that these items would be conditioned by information and estimates regarding total future environments.

III

METHODOLOGICAL MODIFICATIONS

Support for significant modifications and refinements in methodology was provided by (1) specific recommendations on the formal evaluation questionnaire, (2) written comments included in responses to information packages, and (3) encouraging results of specific innovations that were tested in the Sea Grant Delphi exercises. Only the evidence believed to be most significant is covered in this summary.

Safeguards Against a Manipulated Consensus

Several particularly important issues were indicated by panelists in their responses to information packages and evaluation matrixes. The assertion that the Delphi techniques can result in a manipulated and arbitrary consensus resulted in the institution or reemphasis of the following procedures in later rounds of both the technical and broad panel exercises:

1. Items considered by the panel were generally those suggested by the respondents in an earlier round. The source of each suggestion was identified by panel member number, and basic biographical information associated with that number was provided.

2. All pertinent comments, including personal comments, were fed back to the panelists with a minimum of editing.

3. If a respondent missed a round, his estimates for the previous round were carried over--provided that background conditions had not significantly changed.

4. In addition to group statistical summaries, summaries for subgroups--such as the group of individuals who rated their competence relatively high--were part of the information feedback.

5. Supporting arguments were specifically requested from those respondents whose estimates for the previous round were outside of the consensus range. To the extent that the rest of the panelists were influenced by these arguments, the consensus would tend to broaden rather than narrow.

The association of basic biographical information with the source of specific inputs had the additional advantages of assuring the panelists that they were engaged in deliberations with a peer group,[1/] and gave them added guidance in weighing individual

---

1/ For a discussion of circumstances in which complete anonymity could be relaxed, see Murray Turoff, Delphi and Its Potential Impact on Information Systems, Paper 81 (paper presented at the Fall Joint Computer Conference, Nov. 1971).

suggestions and comments. Unfortunately, it is believed that in the technical panel exercises (where the viewpoints and backgrounds of the participants were well known to many of the panelists) the identity of the source of specific inputs was compromised in several instances. In these situations biographical data should be used with great care.

Scaling of Words and Phrases

Personal probability assessments regarding future events and developments are critical in the Delphi Method for obtaining and communicating subjective judgments. It would be highly desirable from several viewpoints if verbal phrases corresponding to commonly used numerical probabilities, such as 10, 25, 50, 75, and 90 per cent, could be identified and widely accepted. The benefits would be the following:

1. Researchers and decision makers would be encouraged to think about a probability scale in similar terms.
2. Verbal phrases are more appropriate than numerical probabilities in expressing the likelihood of future socioeconomic developments. Since this is an area where measurements are less precise than in technological forecasting,[2/] numerical probabilities may give an exaggerated impression of precision.

2/ See Raoul de Brigard and Olaf Helmer. Some Potential Societal Developments--1970-2000, IFF Report R-7 (Middletown, Conn.: Institute for the Future, Apr. 1970).

3. Verbal phrases would minimize the confusion associated with the classical interpretation of probability which limits numerical probabilities to events that can be repeated many times under essentially the same conditions.

Several studies regarding the scaling of words and phrases to express the likelihood of an event have been made,[3/] but there appears to be considerable difference in the results, and several expressions commonly used in the Delphi method are not included in these studies.

In the opening rounds of the Sea Grant Delphi exercises each of the three general participating groups (technicians, behaviorists, and decision makers) was given a list of twenty-three words and phrases commonly used to express the likelihood of an event. Two similar forms were used with one main difference: The words and phrases on Form A were listed in descending order of likelihood--on the basis of the judgment of the researcher, while those on Form B were arranged in a modified random order. Which of the two forms would be sent to each panelist within each group was randomly decided. In each case the respondent was advised that the phrases were not in order, except that those in the left column were generally phrases for a

---

3/ The most comprehensive study appears to be the one done by Sarah Lichtenstein and J. Robert Newman, "Empirical Scaling of Common Verbal Phrases Associated with Numerical Probabilities," Psychonomic Science, IX (Oct. 1967).

better than even chance and those in the right column were generally phrases for a less than even chance. Panelists were asked to assign a numerical probability (ranging from 0 to 100 per cent) that each of the words or phrases brought to mind. The results of the estimates of all respondents are shown in Table 1.

Each group suggested additional words and phrases that they would like to have considered. Estimates of numerical probability for these added items were requested in the following round of information packages. The results for these additional items are shown in Table 2.

Among all the items considered, the words or phrases corresponding closely to the commonly used probabilities of 10, 25, 50, 75, and 90 per cent are shown in Table 3. The median estimate for these expressions of all participants in the Sea Grant Delphi exercises coincides exactly with the commonly used numerical probabilities. The greatest difference in the means is .067 and standard deviations are all less than .103.

Estimates of the numerical probabilities for several expressions that have been frequently used to express intuitively the notion of different degrees or levels of probability showed great dispersion. These are believed by the researcher to be inappropriate for use with the Delphi method. For example, numerical estimates for the word "expected" ranged from .300 to .990, with a standard deviation of .183 and an interquartile range of .325. In addition, the estimates

TABLE 1

Scaled Probability Words and Phrases--Summary for All Respondents

Rank (By Median)	Word or Phrase	Number of Estimates	Mean	Median	Standard Deviation	Interquartile Range Q_1	Interquartile Range Q_3
1	Highly probable	36	0.883	0.900	0.065	0.850	0.900
2	Very probable	36	0.825	0.850	0.085	0.780	0.900
3	Expected	35	0.775	0.800	0.183	0.600	0.925
4	Very likely	36	0.820	0.800	0.077	0.750	0.900
5	Quite likely	36	0.765	0.750	0.102	0.700	0.830
6	3-to-1 odds	36	0.735	0.750	0.065	0.750	0.750
7	Good chance	35	0.712	0.750	0.099	0.600	0.750
8	Likely	36	0.679	0.690	0.100	0.600	0.750
9	Rather likely	36	0.669	0.650	0.095	0.600	0.750
10	Better than even	36	0.583	0.600	0.046	0.550	0.600
11	Slight odds in favor	36	0.561	0.550	0.048	0.550	0.600
12	Toss up	36	0.500	0.500	0.0	0.500	0.500
13	Uncertain	34	0.398	0.490	0.132	0.300	0.500
14	Slight odds against	36	0.433	0.450	0.029	0.400	0.450
15	Somewhat unlikely	36	0.342	0.350	0.086	0.300	0.400
16	Fairly unlikely	36	0.267	0.300	0.092	0.230	0.300
17	Rather unlikely	36	0.241	0.250	0.085	0.200	0.300
18	1-to-3 odds	36	0.258	0.250	0.051	0.250	0.250
19	Not much chance	36	0.147	0.150	0.075	0.100	0.200
20	Quite unlikely	36	0.166	0.110	0.122	0.100	0.250
21	Very unlikely	36	0.129	0.100	0.086	0.050	0.200
22	Improbable	35	0.135	0.100	0.088	0.075	0.180
23	Highly improbable	34	0.070	0.050	0.085	0.020	0.065

TABLE 2

Additional Words and Phrases Suggested and Evaluated by Panelists

Word or Phrase	N†	$\bar{X}$	Median	Standard Deviation	Interquartile Range Q_1	Q_3
Evaluated by Technicians:						
Most likely	10	.765	.750	.091	.700	.800
Probable	10	.665	.650	.067	.600	.675
Very unlikely	10	.121	.100	.056	.100	.150
Earliest possible	10		*	*		
Evaluated by Decision Makers:						
Most likely	10	.796	.850	.088	.665	.850
Probable	10	.646	.600	.096	.575	.680
Possible	10	.488	.500	.181	.350	.625
Evaluated by Behaviorists:						
Most likely	4	.837	.850	.085	.800	.950
Probable	4	.700	.700	.041	.700	.750
Possible	4	.450	.550	.158	.400	.600
Fairly unlikely	4	.262	.250	.025	.250	.300

* Inappropriate.

† Number of estimates.

TABLE 3

Words or Phrases Corresponding Closely to Commonly Used Numerical Probabilities

Commonly Used Numerical Probability (Percentages)	Word or Phrase	Mean	Median	Standard Deviation
10	Very unlikely	.129	.100	.086
	Improbable	.135	.100	.088
25	Rather unlikely	.241	.250	.085
	1-to-3 odds	.258	.250	.051
50	Toss up	.500	.500	.000
75	Good chance	.712	.750	.099
	3-to-1 odds	.735	.750	.065
	Quite likely	.767	.750	.102
90	Highly probable	.833	.900	.065

did not cluster about a single identifiable value. The expression "earliest possible" was evaluated as inappropriate or ignored by all respondents.

It is interesting to note that in cases where the adverbs "quite" and "rather" were attached to "likely" and "unlikely" there was an asymmetry of response. The median response for those expressions with numerical probabilities greater than .50 were about .10 less than one would expect by looking at their mirror images. For example, the median estimates for "quite likely" and "rather likely" were .75 and .65 respectively while the median estimates for "quite unlikely" and "rather unlikely" were .11 and .25. This finding is consistent with those of Lichtenstein and Newman.[4/]

The greatest concern in using numerical probabilities as a means of communication was in the exchange of information between researchers (behaviorists and technicians) and the decision makers. Table 4 provides a comparison between the two groups for expressions that are appropriate for the commonly used probabilities of 10, 25, 50, 75, and 90 per cent. There is a remarkable degree of agreement between the two groups. Where significant differences exist, the estimates of the decision makers are closer to the 50 per cent value than the estimates of the researchers--indicating perhaps a conservative influence.

4/ Ibid.

TABLE 4

Comparison of the Estimates of Sea Grant Researchers ($\bar{X}_R$) and Decision Makers ($\bar{X}_{DM}$)

Commonly Used Numerical Probability (Percentages)	Word or Phrase	$\bar{X}_R$	$\bar{X}_{DM}$	$\bar{X}_R - \bar{X}_{DM}$
10	Very unlikely	.127	.137	.005
	Improbable	.109	.165	-.056
25	Rather unlikely	.205	.281	-.076
	1-to-3 odds	.263	.252	.011
50	Toss up	.500	.500	.000
	Good chance	.711	.714	.003
75	3-to-1 odds	.737	.733	.004
	Quite likely	.800	.726	+.074
90	Highly probable	.909	.854	+.055

A comparison of the estimates of the respondents using Form A with those using Form B (which listed the words and phrases in a different order) indicated that the estimates were influenced very little by changing the order of words. A further check on the reliability of the estimates was conducted by repeating the phrase "very unlikely" on Form B for the technical panel, listing it in both the number 14 and the number 22 positions. The difference in mean estimates was less than .022.

A Delphi methodology was used to determine whether feedback and reassessment would result in greater agreement on the numerical probabilities associated with several expressions. The technique was extremely effective in narrowing the dispersion of the estimates while only slightly changing the measures of central tendency.

Aids to Personal Probability Assessments

Because the technicians had problems in making consistent personal probability assessments, and it was believed that many of the decision makers would have had little experience with the assessment and interpretation of personal probabilities, a guide for making personal estimates of probability was sent to all members of the broad panels.[5/] The guide was based on the concept of a betting rationale to divide the future systematically into four equally attractive

---

5/ The guide for personal assessments of probability is included in the appendix.

segments in estimating the date of occurrence of a future event. The decision makers agreed in their formal evaluation that the guide was helpful, while the behaviorists--who had more experience with probability theory--supported it only mildly as an aid in making personal probability assessments. It is interesting to note that there were no inconsistent probability estimates by the panel of decision makers, while a few inconsistent estimates were made by the behaviorists. The betting rationale technique has additional desirable features for use in Delphi exercises in that it (1) produces individual distributions rather than point estimates, (2) takes advantage of the strong agreement on words and phrases associated with 25, 50, and 75 per cent levels of probability, (3) provides an easily identifiable consensus range, and (4) minimizes the confusion in making estimates for events whose probability the estimator believes will increase for some time and then decrease.

In summary, the results of the research provide strong empirical evidence that there are specific words and phrases denoting the likelihood of a future event which multidisciplinary groups associate with numerical values closely corresponding to the probabilities of 10, 25, 50, and 90 per cent. If used with the numerical values, the phrases could be a factor in making personal probability assessments a more reliable measuring device. The research also provided evidence of a wide dispersion in the numerical values which these same groups assigned to several common expressions used to denote the degree of probability. This latter group of phrases when used

exclusively would be too elastic a measure of personal assessments of probability, and if used in conjunction with numerical probabilities could be a source of ambiguity in a Delphi inquiry.

Relationships Among Forecasted Developments

A criticism of the Delphi method is that it yields a set of independent estimates that do not adequately take into account interdependencies among anticipated developments. Gordon's technique of cross-impact analysis has generally received much support as a technique for adjusting probability estimates to take into account the effect of possible relationships among future events. In the most widely used version of Gordon's technique the estimates made by a Delphi panel are regarded as "initial" estimates which are adjusted using a Monte Carlo approach and many iterations of a rather complex matrix. The matrix is based upon arbitrary quantifications of assumption about the nature of the relationships between events.[6/]

6/ For a more detailed discussion of cross-impact analysis and other techniques for developing internally consistent scenarios see, "The Delphi Method: A Systems Approach to the Utilization of Experts in Technological and Environmental Forecasting," by John D. Ludlow, Working Paper No. 3 (Ann Arbor, Mich.: Bureau of Business Research, Graduate School of Business Administration, University of Michigan, Apr. 1970).

An alternate approach to Gordon's technique which uses the expertise of the Delphi panel and the techniques that are inherent in the basic Delphi method was used in the Sea Grant Delphi exercises. Several procedures were employed to encourage respondents to assess intuitively the relationships among future developments under consideration.

1. Carefully selected multidisciplinary groups were used to obtain different points of view and to increase the probability that relevant developments, believed to be important to the issues, would be suggested and evaluated.
2. Technical experts gave prior consideration to issues particularly vulnerable to technological advances to ensure that requisite technology would not be overlooked.
3. Technical panel members were asked to indicate groups of developments that they believed were closely related. If the occurrence or nonoccurrence of one development would change the probabilities associated with other listed developments they were to consider the developments related.

4. Related issues and developments were considered concurrently.

After making their fourth round judgments regarding all developments, respondents were asked to make conditional probability estimates for developments that panel members had suggested were closely related. One of these groups of developments and the evaluation matrix is shown in Figure 1. The events for this particular series of interrogations will be referred to either as dependent events or as conditioning events. Any single event may be placed in either category as the interactions among pairs of events are sequentially considered.

A sample of an individual response is shown in Table 5. The summary includes the respondent's evaluation of his competence to judge the development--using a scale ranging from 1 (unfamiliar) to 5 (expert or researcher in the area), his initial estimate of the probability of the development occurring in 1971-80, his third round estimate, his final estimate, and his conditional estimates. The computed probabilities that are shown for the dependent events were arrived at using the following formula for combining probability estimates, where A is the dependent event and B is the conditioning event:

$$P(A) = P(B)\ P(A/B) + P(\bar{B})\ P(A/\bar{B})$$

Analysis of all individual responses reveals that a relatively high percentage of respondents altered their final estimates about those developments included in the set of events which was subjected

Round 4

Panel Member 12

Panel T-1

Developments and Events that Respondents Have Suggested Are Interrelated	Probability 1971-80	50% Probability Date
D-32 Requirement by the state, calling for tertiary treatment of municipal sewage for Traverse City		
Your Previous Estimates	80	1977
Panel Estimates, Round 3	75 (50-85)*	1977 (1975-80)
Those Who Rated Competence ≥ 3	83 (62-95)	1978 (1975-80)
Your Next Estimates for D-32		
D-39 Full-scale operation somewhere in the U.S. of physiochemical system for waste water treatment applied directly to primary waste, which is economically competitive with 1970-designed facilities using activated sludge secondary treatments--with chemical precipitation for phosphate removal		
Your Previous Estimates	50	1975
Panel Estimates, Round 3	55 (30-85)	1979 (1975-80)
Those Who Rated Competence ≥ 3	75 (20-90)	1979 (1979-80)
Your Next Estimates for D-39		
D-31 Construction of a spray irrigation system for waste water disposal in the Grand Traverse Bay region		
Your Previous Estimates	50	1980
Panel Estimates, Round 3	50 (50-50)	1980 (1980-90)
Those Who Rated Competence ≥ 3	50 (15-50)	1980 (1978-80)
Your Next Estimates for D-31		
If you were certain that D-32 would occur before 1980, your estimates for D-31 would be		
If you were certain that D-32 would not occur in 1971-80, your estimates for D-31 would be		
If you were certain that D-39 would occur before 1980, your estimates for D-31 would be		
If you were certain that D-39 would not occur in 1971-80, your estimates for D-31 would be		

* Interquartile Range

TABLE 5

Conditional Probability Estimates of Individual Respondent

Devel. No.	Self-Evaluation	Initial Estimate	3rd Round Estimate	Final Estimate	$P(38\mid 80)$	$P(38\mid \overline{80})$	$P(45\mid 38)$	$P(45\mid \overline{38})$	Comp. $P(38)$*	Comp. $P(45)$
80	3	20	20	25						
38	4	60	60	80	100	80			85	
45	3	10	10	10			10	0		8
					$P(84\mid 20)$	$P(84\mid \overline{20})$	$P(7\mid 84)$	$P(7\mid \overline{84})$	Comp. $P(84)$	Comp. $P(7)$
20	3	25	25	35						
84	4	75	75	75	75	75			75	
7	3	50	50	50			100	75		94
					$P(31\mid 32)$	$P(31\mid \overline{32})$	$P(31\mid 39)$	$P(31\mid \overline{39})$	Comp. $P1(31)$	Comp. $P2(31)$
32	3	80	80	75						
39	3	50	50	75						
31	3	50	50	100	100	100	50	100	100	63

* Computed probability of the dependent event using the formula $P(A) = P(B)P(A/B) + P(\overline{B})P(A/\overline{B})$.

to conditional probability assessments. Since this was the third iteration of feedback and reassessment for many of these developments, it is not unreasonable to assume that the change in estimates was primarily due to the evaluation of relationships among events--relationships which previously had not been fully considered. This assumption is further supported by the fact that these respondents made almost no changes in their estimates of other developments in the set which was not subjected to the specific routine of estimating conditional probabilities (but was given the benefit of the feedback of all of the other types of information used in these exercises). Since the relationship among events was stressed throughout these exercises, any movement in the final estimates as a result of the consideration of specific conditioning effects is believed to be significant.

Table 6 summarizes the total effect of the occurrence and nonoccurrence of related events on the probability estimates for those future developments which were subjected to the series of conditional estimates. The table is essentially a sensitivity analysis and tends to confirm that panel members, in the judgment of their multidisciplined contemporaries, were correct in assuming a close relationship among these groups of events.

The medians of the panels' estimates provide a measure of their judgments regarding the sensitivity of the dependent events--measured by the absolute effect on their likelihood of realization during a specific time period (1971-80)--to the occurrence and

TABLE 6

Sensitivity Analysis

Dependent Event (A)		38	45	84	7	31	31
Conditioning Event (B)		80	38	20	84	32	39
Panel Member Number[ff]	47	30*	30	40	30	40	40
	14	21	10	0	25	0	50
	49	50	30	10	20	30	15
	22	25	40	30	10	25	4
	44	45	5	-	8	5	5
	40	35	50	55	20	-	-
	18	30	20	0	0	40	80
	32	-	-	-	-	25	-
	2	25	0	0	0	-	-
	38	50	5	25	40	20	20
		75	40	50	10	50	70
Median		32.5	25.0	25.0	15.0	25.0	30.0
$Q_3 - Q_1$		37.5	22.5	40.0	8.5	20.0	50.0

* The numbers in the double-ruled area are the absolute change in the probability of the dependent event computed using the formula $|P(A/B) - P(A/\bar{B})|$.

ff A factor has been applied to actual panel member numbers to obtain the numbers used in this figure.

nonoccurrence of a conditioning event. The interquartile ranges provide a measure of the uncertainty among panel members about the nature of the relationships between the pairs of events.

This type of analysis provides a decision maker with information about actions which offer the greatest potential for influencing a development of particular interest to him. It also offers insight into the uncertainty among panel members concerning the secondary effects of conditioning developments. For example, the related developments of Group 3 (Table 5) present two types of conditioning events. One event (39) was the demonstration of waste water treatment methods which incorporated significant advances in technology, while the other event (32) dealt with political intervention in the form of legal water-quality standards. The group's judgment was that either conditioning event would have considerable influence on the construction of a spray irrigation system for waste water disposal in the Grand Traverse Bay region. The median estimates for the absolute change in the probability of the dependent event (31) were 25 and 30 per cent respectively, considering the occurrence and nonoccurrence of conditioning events 32 and 39 (Table 6). However, the group was more uncertain about the conditioning effects of the demonstration of an advanced treatment method than they were about the conditioning effects of a specific form of political intervention--the interquartile ranges for the group estimates being 50 and 20 per cent respectively.

Self-evaluation indexes based on competence in the area under

consideration and confidence indexes associated with conditional probability estimates offer a mechanism for weighing individual estimates and intuitively improving the credibility of group judgments. For example, a decision maker may have estimates by experts who would feel quite confident in making probability estimates of a development (Development A) occurring in a given time period if they could be sure of the occurrence or nonoccurrence of a related event (Development B)--but these experts do not feel competent to evaluate the related event. If other experts are available to the decision maker who are competent in evaluating the conditioning event, then combining the average estimates of the first group regarding $P(A/B)$ and $P(A/\bar{B})$ with those of the second group regarding $P(B)$, using the formula $P(A) = P(B)\ P(A/B) + P(\bar{B})\ P(A/\bar{B})$, provides him with estimates that are intuitively better than independent estimates of A.

In making their estimates regarding important technical developments panelists were asked to assume a future social and political environment which was consistent with present trends.[7/] It is highly unlikely that all panelists made similar assumptions, especially in view of the wide range of training and experience they represented.

7/ This approach was suggested in Selwyn Enzer, "A Case Study Using Forecasting as a Decision-making Aid," IFF Working Paper, WP-2 (Middletown, Conn.: Institute for the Future, Dec. 1969).

For example, panelists may have discounted at different rates the effects of political actions that they felt would have a strong conditioning effect on a technical development because these actions were not consistent with present trends--even though they felt the conditioning political actions had a high probability of occurring during the period under consideration. Requesting estimates which were conditional on various types of political intervention would provide insight as to the individual estimator's assumptions about the future political environment.

In summary, our experience in the Sea Grant Delphi exercises with systematically obtaining and analyzing conditional probability estimates demonstrated the feasibility of the technique for:

1. Encouraging respondents to re-examine their estimates, taking into consideration the influence of related developments
2. Assessing the extent to which relationships among anticipated developments were intuitively accounted for by panelists in the series of estimates and information feedback that are inherent in the Delphi method
3. Identifying the influence of conditioning events and the uncertainty associated with their effect
4. Assessing the effectiveness of alternative strategies, such as support for specific research efforts, that will most effectively influence the outcome of important developments

5. Providing, in combination with self-evaluation and confidence indexes, a mechanism for weighing individual estimates and intuitively improving the credibility of group judgments
6. Determining some of the respondents' assumptions regarding the total social, political, and technical environment when they made judgments about specific developments

IV

EVALUATION OF THE DELPHI METHOD

There is far from universal agreement on the merits of the Delphi techniques. RAND believes that Delphi marks the beginning of a whole new field of research, which it labels "opinion technology."[1/] However, a paper presented at the joint statistical meetings of the American Statistical Association in August 1971 described the Delphi techniques as the antithesis of scientific forecasting and of questionable practical credibility.[2/]

According to a recent Wall Street Journal article the Delphi technique is gaining rather widespread use in technological forecasting and corporate planning, although the same article cautions:

> It's easy enough to see the shortcomings of the Delphi procedure; it's much harder to rectify them, as many are struggling to do. Remedial work must be done if the method is to be used in good conscience.[3/]

1/ "Forecasters Turn to Group Guesswork," Business Week, Mar. 14, 1970.

2/ Gordon A. Welty, "A Critique of the Delphi Technique" (summary of paper presented at the Joint Statistical Meetings of the American Statistical Association, Colorado State University, Fort Collins, Colorado, Aug. 23-26, 1971).

3/ "Futuriasis: Epidemic of the '70s," Wall Street Journal, May 11, 1971.

The Sea Grant Delphi exercises offered an exceptional opportunity for a critical evaluation of the Delphi techniques in an operational environment. The panelists--the main resource in evaluating the methodology--were interested in the improvement of techniques to integrate the judgments of a multidisciplinary research team and to convey its informed insights to society. Their evaluations were not biased by a strong emotional involvement in the success of the exercises as has been true with many of the individual assessments of the method that have been published. From both a program budgeting standpoint and demands on researchers' time, the Delphi exercises competed with a wide variety of other methods for securing and disseminating information.

The primary instrument in evaluating the effectiveness of the Delphi technique in the Sea Grant exercises and its potential in other applications was a formal questionnaire (see Appendix I). This was developed almost entirely by the respondents themselves using the Delphi technique of feeding back collated individual suggestions to generate additional suggestions. A scaled descriptor was then used to quantify the subjective judgments of panelists regarding the suggested items. This procedure somewhat reduces the questionnaire's vulnerability to the biases and shortcomings of the investigator. To supplement the formal questionnaire, over thirty-five interviews with panelists were conducted.

The third-round information package for the technical panel included a form for obtaining the suggestions of panelists about positive and negative aspects of the Delphi techniques and appropriate areas for applications of the methodology. One example was provided by the administrator for each category of response desired. For example:

Positive aspect:	It permits an orderly development of a well-considered opinion.
Negative aspect:	It is superficial.
Suggested application:	To develop criteria for recruiting industries to a region, considering their desirability, feasibility, and impact on water resources.

The collated responses yielded thirty-one discrete suggestions which included seven positive aspects of the techniques, thirteen negative aspects, five appropriate areas of applications, and six modifications. In the evaluation forms of the subsequent round, three questions about Sea Grant objectives and one question to determine the respondent's familiarity with the Delphi method prior to the Sea Grant exercises were added.

The information package also contained two recent articles on the Delphi method. An article by Murray Turoff, "The Delphi Conference,"[4/] was included because it discussed a concept similar

---

4/ Murray Turoff, "The Delphi Conference," The Futurist, Apr. 1971.

to one being considered for use in conjunction with Sea Grant seminars and planning meetings. A Wall Street Journal article of May 11, 1971, "Futuriasis: Epidemic of the 70 s," was included because it criticized the Delphi method and thus was expected to offset some of the favorable comments of the researcher in his letters to the panel soliciting their participation in the exercises. The articles were designed to encourage a more objective evaluation of the methodology.

The evaluation questionnaire submitted to members of the broad panels was similar to the one considered by the technical panel with these exceptions:

1. All items suggested by the researcher were eliminated so as to remove a potential source of bias.
2. Items dealing with modifications of the Delphi procedures (six items) were removed because either they were believed to be more appropriate for the technical panel or the ideas were incorporated in other statements.
3. An item was added to evaluate the usefulness of the betting rationale approach in personal probability assessments--used only with the broad panels.
4. An item was added to determine if broad panel respondents would consider using the Delphi method if they were responsible for decisions involving regional planning and development.

The evaluation items were in the form of assertive statements with which the respondent was asked to agree or disagree using a six-point scale. Panelists were to evaluate each statement from two points of reference: first, their experience in the Sea Grant Delphi exercises, and second, their expectations of the method's potential applications which incorporated their suggested modifications and which would be administered by a professional team of researchers and administrators.

The primary interest in this research is the evaluation of the Delphi method's potential, since the Sea Grant exercises expose the panelists not only to the techniques as they have been usually employed but also to essentially new applications and modifications, some of which may be judged inappropriate for the present exercises and yet be useful in other situations. A comparison of how closely the present exercises correspond to the panelists' conception of an ideal treatment provided information that was helpful in analyzing the judgments of the panelists and in designing subsequent Sea Grant Delphi exercises. In this aspect of the evaluation the judgments of panelists who had previous experience with the Delphi method were of particular interest.

Summaries of the evaluation indicate the proportion of numerical values assigned by the respondents as well as the number of estimates (n), a measure of central tendency ($\bar{x}$), and a measure of dispersion (σ). The descriptive words associated with the six-point scale used to measure degrees of agreement or disagreement were:

strongly disagree = 1	somewhat agree = 4
disagree = 2	agree = 5
somewhat disagree = 3	strongly agree = 6

In recording their judgments several panel members circled two adjacent numbers and in these cases the number that prevailed was randomly decided. A value of 3.5 was viewed as the neutral point in interpreting the group means. The boundaries for the descriptive words and phrases are shown below:

strongly disagree	disagree	somewhat disagree	somewhat agree	agree	strongly agree
1	2	3	4	5	6
1.5		2.5	3.5	4.5	5.5

Summaries are provided for the three general groups participating in the Sea Grant Delphi exercises: technicians (Group I), behaviorists (Group II), and decision makers (Group III). For some issues the summaries for technical panelists under 40 years of age and panelists with previous experience with the Delphi method are shown. Using the sample results, tests of significance were made to test the hypothesis that the distributions of the judgments of the Delphi method are homogeneous across the groups (the test procedure was based on the chi-square test statistic) and to test the null hypothesis that the means of the judgments of the population represented by the groups are identical (based on analysis of variance and the F-test). The results of these tests are used to support the discovery of basic differences in judgments

made by different groups which have been formed on the basis of similar backgrounds and experiences. Where appropriate, the probability of obtaining a value larger than the observed statistic (chi-square or F) is indicated.

Some insight into the nature of the difference between judgments based on panelists' experience in the Sea Grant Delphi exercises and the panelists' conception of an ideal application of the Delphi techniques can be gained by examining Table 8, which is a cumulative summary of the evaluation of the effectiveness of the method in three specific roles. One role--obtaining, combining, and displaying the opinion of informed people--is the role the method was designed for. The other two roles, to encourage greater involvement and to facilitate communication between researchers and decision makers, are essentially new roles for the method--at least in terms of emphasis.

The greatest difference in judgments based on experience and potential is with the behaviorists' group ($\overline{X}_e = 4.3$, $\overline{X}_p = 5.1$), and the least is with the decision makers ($\overline{X}_e = 5.0$, $\overline{X}_p = 5.2$). On the basis of the average judgments for all respondents, the Sea Grant Delphi exercises corresponded rather closely to the panelists' conception of an ideal treatment ($\overline{X}_e = 4.2$, $\overline{X}_p = 4.6$), and there is a similar spread in the judgments of the subgroup with previous experience with the method ($\overline{X}_e = 4.5$, $\overline{X}_p = 4.9$).

TABLE 7

Comparison of Evaluations Based on Experience in Sea Grant Exercises with Evaluations Based on the Potential of the Method

Group			Distribution of Judgments (proportion of total)								
			1	2	3	4	5	6	n	$\bar{x}$	σ
I	Technicians	Experience	.11	.14	.08	.44	.19	.03	36	3.6	1.340
		Potential	.05	.05	.16	.39	.30	.05	57	4.0	1.172
II	Behaviorists	Experience	.00	.00	.10	.65	.15	.10	20	4.3	.696
		Potential	.00	.00	.04	.15	.42	.32	24	5.1	.859
III	Decision Makers	Experience	.00	.00	.00	.33	.37	.30	30	5.0	.809
		Potential	.00	.00	.00	.13	.50	.37	30	5.2	.679
I, II,& III	All Respondents	Experience	.05	.06	.06	.45	.24	.14	83	4.2	1.218
		Potential	.03	.03	.09	.27	.40	.19	111	4.6	1.101
	Respondents under 40	Experience	.07	.02	.07	.54	.22	.07	41	4.0	1.172
		Potential	.05	.00	.09	.21	.45	.20	56	4.6	1.218
	Previous Delphi experience	Experience	.00	.00	.08	.49	.31	.13	39	4.5	.941
		Potential	.00	.00	.11	.17	.45	.28	47	4.9	.938

Effectiveness in Specific Functions

The statistics related to the potential of the Delphi method are of greatest interest in evaluating the effectiveness of the method in specific functions. There appears to be a relatively small and stable difference between the evaluations based on experience in the Sea Grant exercises and those based on an ideal application of the method. Also, any future Sea Grant Delphi exercises will incorporate the recommendations of these panelists, many of whom are experts in methodology.

The evaluation of the effectiveness of the Delphi method in obtaining, combining, and displaying the opinion of informed people indicated that the technicians on the average agree somewhat that it was effective compared with alternative methods ($\overline{X}_T = 4.0$). There is, however, considerable dispersion in their estimates ($\sigma = 1.172$) with some respondents strongly disagreeing that the method was effective in this role. The behaviorists agree that it was effective in this designed role ($\overline{X}_B = 4.7$), and the decision makers displayed strong agreement in this area ($\overline{X}_{DM} = 5.3$). The dispersion in the estimates of these two latter groups is much less than was the case with the technicians. An analysis of variance on these data gave a value of 4.6233 for the F statistic. The probability of obtaining a larger F value than 4.6233 when the groups are identical is .0165. A chi-square analysis gives a value for chi-square of 8.3588. The probability of obtaining a larger value, when in fact the distributions come from a homogeneous population, is .2130--the descriptive level of significance. A combination of the judgments of the decision makers and the behaviorists resulted in $P(\chi^2 > 9.165) = .0025$ (see Table 9).

TABLE 8

Effectiveness in Obtaining, Combining, and Displaying the Opinion of Informed People

Group			Distribution of Judgments (proportion of total)								
			1	2	3	4	5	6	n	$\bar{x}$	σ
I	Technicians	Experience	.11	.11	.06	.50	.22	.00	18	3.6	1.290
		Potential	.05	.05	.11	.42	.32	.05	19	4.1	1.177
II	Behaviorists	Experience	.00	.00	.00	.72	.29	.00	7	4.3	.488
		Potential	.00	.00	.00	.43	.43	.14	7	4.7	.756
III	Decision Makers	Experience	.00	.00	.00	.22	.22	.56	9	5.3	.675
		Potential	.00	.00	.00	.10	.50	.40	10	5.3	.675
I, II,& III	All Respondents	Experience	.06	.06	.03	.47	.23	.15	34	4.2	1.274
		Potential	.03	.03	.06	.33	.39	.17	36	4.5	1.108
	Respondents under 40	Experience	.06	.00	.00	.69	.19	.06	16	4.1	1.025
		Potential	.05	.00	.00	.21	.58	.16	19	4.7	1.098
	Previous Delphi experience	Experience	.00	.07	.00	.47	.27	.20	15	4.5	1.060
		Potential	.00	.00	.06	.19	.50	.25	16	4.9	.085

Combining Groups II and III:

$P(F > 4.6233) = .0165$

$P(\chi_2^2 > 8.3588) = .2130$

$P(\chi^2 > 3.959) = .0464$

The evaluation of the method's effectiveness in encouraging greater involvement in Sea Grant activities provided similar results. The average judgments of effectiveness were $\overline{X}_T = 4.0$, $\overline{X}_B = 4.7$, and $\overline{X}_{DM} = 5.1$, and the dispersion in the estimates of the technicians was much greater than for the other two groups. The descriptive levels of significance are .0026 based on the F distribution, and .0099 based on the chi-square distribution.

The average judgments regarding the method's effectiveness in communicating information to regional planners and decision makers were $\overline{X}_T = 4.1$, $\overline{X}_B = 5.4$, and $\overline{X}_{DM} = 5.3$. The judgments of the behaviorists were exceptionally high regarding this use of the method--essentially a new role in terms of emphasis--and may reflect this group's special concern for the psychological and sociological barriers associated with alternate methods of obtaining group judgments. The descriptive level of significance is .0003 based on the F distribution, and .0108 based on the chi-square distribution.

For all three roles there appears to be little difference in the average estimates aggregated according to the respondent's age and the average estimates of all respondents. However, respondents with previous Delphi experience showed substantially higher average estimates concerning the method's effectiveness in obtaining, combining, and displaying subjective judgments and in encouraging involvement, but had approximately the same average estimates about the communication role as the other respondents.

Effectiveness of the Method in Specific Situations

Table 9 provides statistics of the panelists' judgments regarding specific applications for the Delphi method that were suggested by the respondents and also the recommendations of the broad panelists about the use of the Delphi method as an aid in decision making. The average judgments of all respondents reflect support for the use of the Delphi method in two areas: (1) developing criteria to recruit industries in a region--considering their desirability, feasibility, and impact on water resources ($\overline{X} = 4.6$), and (2) employing it for situations dealing with the future, uncertainty, and conflicting views ($\overline{X} = 4.5$). The groups' judgments are affirmative about establishing program priorities within a program committee ($\overline{X} = 4.4$), long-range planning by a university ($\overline{X} = 4.1$), the education of practicing politicians by including them in the exercises ($\overline{X} = 4.3$), and budgeting of an interdisciplinary program ($\overline{X} = 4.1$). The support in this latter group of specific applications is generally strongest among the decision makers and weakest among the technicians. Only the behaviorists and decision makers evaluated the method as an aid in decision making, and both groups supported its use in this role ($\overline{X}_B = 5.1$, $\overline{X}_{DM} = 5.0$, $\sigma_B = .835$, $\sigma_{DM} = 1.054$).

Tailoring the Method to Groups

The evaluation results showed that among the carefully selected samples of people in the Sea Grant exercises the techniques were more highly regarded by multidisciplinary groups than by technicians.

TABLE 9

Suitable Applications for a Delphi Methodology

Group		Develop Industry Criteria		Budgeting		Education	
		$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ
I	Technicians	4.3	.996	3.9	1.356	4.1	.793
II	Behaviorists	4.5	1.130	3.7	.866	4.3	.707
III	Decision Makers	5.3	.675	4.8	.789	4.6	1.424
	All Respondents	4.6	1.081	4.1	1.164	4.3	.954

Continued

TABLE 9

Suitable Applications for a Delphi Methodology--Continued

	Group	Uncertainty Future Conflict		Long Range Planning		Program Priorities		Decision Making Aid	
		x	σ	x	σ	x	σ	x	σ
I	Technicians	4.5	.943	3.8	1.393	3.9	1.393	--	--
II	Behaviorists	4.3	1.165	4.1	1.453	4.6	1.453	5.1	.835
III	Decision Makers	4.5	.707	4.8	1.080	5.0	1.080	5.0	1.054
	All Respondents	4.5	.919	4.1	1.355	4.4	1.355	5.1	.938

which is the group most designers and administrators of the techniques have focused on. Of greater significance, however, is the fact that the method can be tailored to the needs and expectations of a group on the basis of its background, training, and experience.

In evaluating positive aspects of the Delphi method that had been suggested by panelists, all groups agree that it offers "the advantages of a committee approach without the necessity of meeting." The behaviorists and decision makers agree that it "can get experts involved because preparatory work is done for them and they need to contribute only numerical estimates or brief statements," but the technicians are polarized in their judgments of this suggested point. The decision makers agree that its ability to provide "a better comprehension of the larger picture" is a positive aspect of the method, while the technicians are split in their judgments and the behaviorists see this as a slightly positive aspect. The behaviorists and decision makers strongly support the statement that "the techniques give one a broad perspective of potential technical developments related to a particular area," while the technicians only somewhat agree. Two parts of the method which were strongly agreed to be positive aspects by all groups are: (1) Introducing ideas not likely to be considered by any one individual, and (2) systematically refining issues as the exercise progresses. To the technicians one of the most attractive aspects of the method is that its procedures are less susceptible to influence by a strong personality than are other procedures. This is also strongly supported as a positive aspect by the other two groups.

Three additional aspects suggested as positive qualities and judged only by the technical panel were: an orderly development of a well-considered opinion, flexibility to expand or narrow the information exchange in accordance with the nature of the issue, and participation at one's convenience rather than at a specific time.

Many of the method's aspects that were judged to be negative were related to desirable aspects of conferences that are not realized with the Delphi method. These included the procedure's failure to provide any opportunity to: (1) change and exchange opinions, (2) hear reasons for various opinions, (3) benefit from the cross-fertilization of the others' ideas, sometimes better informed, and (4) clear up any ambiguities in terms and questions. Interviews with panelists confirm that the Delphi method will be more acceptable when conventional techniques for securing group judgments--such as personal interviews, committee meetings, and seminars--are integrated with the basic Delphi procedures.

All groups rejected the suggestion that the Delphi method was not appropriate for complex, interdisciplinary problems--especially the behaviorists. The discovery of symmetry in the distribution of judgments regarding pairs of items that are similar (although one is suggested as a positive and one a negative aspect) indicates the reliability of the panelists' judgments in the Sea Grant Delphi exercises. The suggestion that the time-consuming aspects of the method are a negative aspect was judged near the neutral position for all three

groups. But there was a large dispersion in the overall judgments for the three groups. This finding is especially interesting since it is believed that the Sea Grant Delphi exercises made greater demands on panelists' time than could be expected in most Delphi applications. In addition, the Sea Grant Delphi exercises competed with the panelists' regular obligations--in contrast to most Delphi exercises in which participation is expected as part of a job or has some reciprocity implications. The suggestion that the routine "tends to be tedious" received a similar neutral judgment from all groups except the technicians, who for the most part agreed that it was a negative aspect. Interviews with technical panelists revealed that some who were specialists in well-defined areas had very little interest in many of the issues and developments being considered but felt obligated--from the standpoint of being a bona fide panel member--to formulate judgments on most of the items.

The suggestion that the Delphi exercises "lack any relation to the economic realities of the times" was solidly rejected as a negative aspect by all three groups. The suggestion that "one's judgment is frequently biased by constraints of the time" received on the average a neutral judgment.

There was slight agreement that the statements "very uninformed judgments are given to many questions," "superficial," and "it should provide an overview but it seems rather like a guessing game" represent negative aspects of the method. However, contrary to the generally

neutral positions indicated on the questionnaire, in interviews with panelists the following criticisms of the method were mentioned more than any others. The statements included, "I feel I am a junior member of a very talented group"; "I would like more facts and more research carried out before expressing a viewpoint"; "I find it easier to make estimates regarding items I know very little about than making estimates in areas I am supposed to be familiar with " and "I did not feel that in the exercise I made any contribution beyond what an interested individual in the region could have made." This suggests that panel members, and technicians especially, must be persuaded that intuitive judgments are valuable even if not based on the quality of research information and quantitative data which they are accustomed to seeing before making a judgment.

The panel agreed that there should be more emphasis placed on the idea that an expert should "not feel obligated to express an opinion on every issue." However the Sea Grant Delphi exercises stressed the concepts of a systems approach and multidisciplinary teams. Therefore it was desired that each respondent consider all items and attempt estimates on those with which he had some familiarity. A self-evaluation index was provided so that a panelist could assess his competence regarding each item. The competence index was a control factor in developing the statistical summaries that were part of the information feedback. This procedure allows an informed person to evaluate relative importance and desirability--an evaluation which

he can do without being an expert in the area--and gives the administrator additional assurance that panelists considered items outside of their specialized areas. In addition, there was some interest in comparing the estimates of experts and nonexperts on specific issues.

There was great dispersion in judgments regarding "experts interfacing with a computer in a seminar or gaming situation," and "the estimates for each round should require twenty minutes or less of a respondent's time." The average judgment of the latter statement ($\overline{X}_T = 4.1$) is consistent with the technical group's judgment of the statement "tends to be tedious" ($\overline{X}_T = 4.4$).

There was also great dispersion in the judgments of the statement that the method would be "better employed after a panel of experts has formulated a tentative report of recommendations" and in the statement that "judgments sought may have been structured with a presentation of known facts in many matters at the start of the study." A portion of this disagreement can be explained by the fact that the panel membership is almost equally divided between Sea Grant researchers who have been considering many of these issues for over a year and experts who have no association with the Sea Grant Program.

The suggestion that the method "can result in a manipulated and arbitrary consensus" received a neutral judgment from all three groups, perhaps reflecting that the respondents felt this danger is no greater than in alternative techniques for securing group judgments. However, on the basis of my experience as a Delphi administrator

it is this researcher's opinion that the Delphi techniques could be a powerful tool for manipulating opinion and policy. For example, as suggested and endorsed by the panelists in these exercises, a suitable application of the method would be the "education of practicing politicians by being included in Delphi exercises." In the Sea Grant exercises panelists were asked to make estimates about the trends of social, political, and economic indicators which have been generally used to gauge the social and economic development of a region. The primary purpose of this aspect of the Delphi exercises was to provide reasonable reference points or boundaries which would temper subsequent highly subjective evaluations. It would appear that the Delphi method could be easily transformed from an educating and guidance function to one of manipulation.[5/]

Statements made by panelists during interviews also reflected a concern about an engineered consensus. Some examples were: "If the sample consists of program committee people primarily, there will be a great deal of similarity in the responses"; "the results are biased

---

5/ For a discussion of the dangers associated with a Delphi devoted to policy issues see "The Design of a Policy Delphi," by Murray Turoff, Technological Forecasting and Social Change, II (Nov. 1970). For a discussion of deliberate distortion see "A Critique of Some Long-Range Forecasting Developments," by Gordon Welty (paper presented at 38th session of the International Statistical Institute, Washington, D.C., Aug. 1971).

by the form and content of the questions"; "problems may indeed be opportunities, but by listing them under problems you bias the whole exercise"; and "instead of bringing in new ideas you tend to perpetuate the old ways of doing things." Some of the procedures used in the Sea Grant Delphi exercises to make a manipulated consensus less likely have already been discussed in the section on methodological modifications. While these measures perhaps go beyond precautions taken in other Delphi exercises, they may not be adequate (in the opinion of the researcher) if the designers and administrators of the exercises desire specific results which may support their own opinions or goals.

Specific modifications of the Delphi method were suggested by the technical panel. There was solid support for "more provisions for constrained judgments--that is, an 'if this, then that' provision." One approach that was used in the Sea Grant exercises--utilizing sequential conditional probability estimates--appears to have considerable potential for providing this capability. This approach is discussed under "Relationships among Forecasted Developments" (Chapter III).

Technical panel respondents also agreed that the feature "safeguards for foreclosing future options" is desirable and should be incorporated. It was suggested that this could best be accomplished if the Delphi procedures were programmed for use with computers so that a respondent would have access to the entire data bank and statistical summaries of previous estimates, thus allowing him to reassess his judgments at any time.

Some of the negative assessments of the method or the suggested modifications can be related to the panelists' previous experience with conventional surveys and questionnaires which emphasize short and quick responses from a randomly selected segment of a population. These suggested modifications included: (1) "The method should require twenty minutes or less of a respondent's time," (2) "It is better employed after a panel of experts have formulated a tentative report of recommendations," and (3) "The statistics provided by relatively uninformed people mean nothing."

In order to minimize the association of Delphi interrogations with conventional techniques, the use of the word questionnaire was kept to a minimum, substituting such expressions as information packages and evaluation matrixes. The progressive or cumulative design of exercises was also emphasized--starting with a brainstorming concept in the early rounds and culminating in carefully formed judgments on issues that had been declared important, within the competence of the panel, and in which areas of disagreement existed. A careful selection of panelists was indicated and basic biographical data on the panel were available to all respondents and in many cases were associated with specific inputs. Scaled self-evaluations of personal competence were also available. Some indication of the success of these measures can be drawn from the fact that only the suggestion "should require twenty minutes or less of a respondent's time" received any support ($\overline{X} = 4.1$). This item was evaluated only by the technicians and there was considerable dispersion in their estimates ($\sigma = 1.409$).

The Degree of Participation

Tables 10 through 12 reveal the extent of each panel member's participation in the Sea Grant Delphi exercises as measured by written suggestions and estimates for each of the four rounds, completion of a formal evaluation of the method, and participation in an interview with the researcher. It is interesting to note that participation by University researchers (technicians and behaviorists) was stronger on the fourth round than on the first round. Nine panelists left the University prior to the final round so that 30 of 49 potential respondents submitted written evaluations on the first round and 29 of 40 potential candidates submitted evaluations on the final round. The total number of panelists who participated in the evaluation of the methodology (by written evaluation or interview) was forty-one. The number of decision makers submitting written evaluations was generally stable after the first round. Sixteen of twenty-two potential candidates participated in the evaluation of the methodology. Since the decision makers were scattered throughout a ten-county area, it is not known how many of the seven who did not respond were not available. Among the University researchers, the participation of those not associated with the Sea Grant program was at least as high--as measured by written responses--as of those directly involved in the Sea Grant program.

TABLE 10

Participation in Sea Grant Delphi Exercises--Technical Panel

Member Number	Written Evaluations of Materials Round 1	2	3	4	Written Evaluation of Methodology	Interview	Unavailable After
026*				x	x		
027						x	
028	x			x	x	x	
029	x	x	x	x	x	x	
030		x	x			x	
022			x		x	x	Round 3
024				x	x	x	
036	x	x	x	x	x		
037	x		x	x	x	x	
038	x		x	x	x	x	
039						x	
040	x			x	x	x	
041				x		x	
042	x	x	x	x	x		
043	x	x	x	x	x	x	
044		x					
045		x					Round 1

* A factor has been applied to actual panel member numbers to obtain the numbers used in this table.

Continued

TABLE 10--(Continued)

Participation in Sea Grant Delphi Exercises--Technical Panel

Member Number	Written Evaluations of Materials Round 1	2	3	4	Written Evaluation of Methodology	Interview	Unavailable After
047						x	
048	x						Round 2
049	x					x	Round 2
050	x	x	x	x	x	x	
051	x	x	x	x	x	x	
053	x		x	x	x		
054	x	x	x	x	x	x	
055	x	x					Round 2
056	x		x	x	x	x	
057		x		x	x	x	
059						x	
060	x	x	x	x	x		
061						x	
062	x	x		x	x		
063				x	x	x	
064	x	x				x	Round 3
Group Totals 33	20	14	14	20	20	23	6

TABLE 11

Participation in Sea Grant Delphi Exercises--Behaviorists

Member Number	Written Evaluations of Materials Round 1	2	3	4	Written Evaluation of Methodology	Interview
126*				x	x	x
128						x
129	x	x	x	x	x	x
130						
134	x			x	x	x
136						
137						
139	x	x	x	x	x	
140	x					x
141	x	x	x	x	x	
142	x		x	x	x	x
146	x	x	x	x	x	x
147	x			x	x	x
148	x					x
150	x		x	x	x	x
151						
Group Totals 16	10	4	6	9	9	9

* A factor has been applied to actual panel member numbers to obtain the numbers used in this table.

TABLE 12

Participation in Sea Grant Delphi Exercises--Decision Makers

Member Number	Written Evaluations of Materials Round 1	2	3	4	Written Evaluation of Methodology	Interview	Unavailable After
076*	x						
077	x	x	x	x	x	x	
078	x						
079	x						
080	x					x	
081	x		x			x	Round 3
082	x					x	Round 2
083	x	x		x	x	x	
084							
085	x	x	x	x	x	x	
086	x						Round 2
087	x	x	x	x	x		
088	x					x	
089	x	x	x	x	x		
090	x						
091	x	x	x	x	x	x	
092	x	x	x	x	x		
093	x	x	x	x	x	x	
094	x					x	
095	x	x	x	x		x	
096		x		x	x	x	
097		x		x	x	x	
Group Totals 22	21	11	9	11	10	12	3

* A factor has been applied to actual panel member numbers to obtain the numbers

V

SUMMARY OF FINDINGS AND RECOMMENDATIONS

The Sea Grant Delphi exercises were designed to obtain and refine the judgments of an interdisciplinary group of researchers. The judgments are about issues and developments that should be considered when planning for intelligent management of the water resources of the Great Lakes.

An important objective of the exercises was to convey the judgments of the researchers to the communities which are to benefit from the research. One approach toward this objective was to include on the panels--on the same basis as the researchers--people who were believed to be influential in the political processes through which regional planning is accomplished. Their knowledge of the issues and the region was not only beneficial to the deliberations, but more importantly their participation was judged to be an effective way of communicating information to regional planners and decision makers.

Two of the three panels were made up of researchers who were designated, on the basis of their training and experience, as technicians and behaviorists. The third group was made up of concerned

citizens who were designated as decision makers. In addition to forecasting, the method was used in several other roles involving the quantification of subjective judgments. The exercises were designed to be progressive and cumulative, with an emphasis on an orderly development of informed judgments.

Some innovations were introduced and others suggested to minimize aspects of the techniques which had met with negative assessments. These concerned the threat of a manipulated consensus, the desire for constrained or conditional judgments, and the desirable aspects of interpersonal methods not available using the Delphi techniques exclusively.

Special attention was given to the communication aspect of the method so that information could be most effectively transmitted to the decision makers. Special attention was also given to improving the reliability of personal assessments of probability. This was accomplished through the identification and scaling of expressions that correspond closely with commonly used numerical probabilities and through the use of a betting rationale to divide the future systematically into equally attractive segments.

The feasibility of incorporating personal probability assessments which consider both the occurrence and nonoccurrence of conditioning events was demonstrated. This technique, in combination with other procedures designed to encourage respondents to assess intuitively relationships among future developments, utilizes the expertise and

processes inherent in the basic Delphi method. In addition, it offers a method of assessing the effect of alternate strategies and random events as well as a mechanism for weighing and intuitively improving group judgments.

The primary resource in evaluating the methodology was the respondents in these exercises--a group of people with exceptional qualifications representing a broad range of academic disciplines and experience. The panelists were interested in developing improved techniques to convey to society informed insights about the future as well as the substantive results of the exercises.

The evaluations showed that among the carefully selected samples of people the techniques were more highly regarded by groups formed on the basis of broad ranges in training and experience than by technicians--the group most administrators of the techniques have focused on. Evaluation results also support tailoring the method to groups on the basis of their background training and experience.

There was agreement that the method was effective in obtaining, combining, and displaying the opinion of informed people--the declared function of the techniques. The method was also evaluated as equally or more effective in two related management tasks that, in terms of emphasis, are essentially new applications for the methodology: (1) encouraging greater involvement, and (2) effecting meaningful communication between researchers and decision makers.

This Delphi inquiry supports a wider market for the method than it presently has by: (1) Identifying the technique's features

that are perceived as desirable and undesirable by groups which are formed on the basis of their training and experience, (2) employing the techniques in essentially new functions and roles, and (3) testing several significant modifications and refinements.

Recommendations for Future Applications of the Method

The design and administration of a Delphi exercise in which the concept of a multidisciplinary team and a systems approach is desired can best be handled as a project within a professional research organization.[1/] The scope of the exercise is generally determined by the respondents and as interesting and unexpected issues are suggested flexibility is needed in designing evaluation matrixes and in determining the composition of the panels. Experts knowledgeable in specialized areas should be available on an ad hoc basis to formulate questions and collate responses in order to minimize redundancy and ambiguity. The demand for their services in the course of a Delphi exercise is very uneven, as is the need for designing, editing, typing, and distribution services. There are significant start-up and learning

---

1/ Murray Turoff's article on the design of a policy Delphi is recommended for anyone interested in designing or participating in a Delphi exercise. See "The Design of a Policy Delphi," Technological Forecasting and Social Change, II (Nov. 1970).

costs associated with the Delphi techniques that can be justified only if the technique will become a routine management tool to be used on a continuing basis. This is particularly true if the benefits of computer processing are to be realized.

The following are some general observations that are consistent with the items suggested and evaluated by respondents in the Sea Grant exercises and with the information gained in personal interviews with panelists.

1. Respondents will be more receptive if the techniques are tailored to specific groups on the basis of their training and experience.
2. The administrator should consistently emphasize the distinction between the characteristics of a Delphi interrogation and those of conventional questionnaires and polls.
3. Panelists--particularly those with technical backgrounds--must be convinced that judgments often have to be made about issues before all facets of the problems have been researched and analyzed to the extent they would like. (For these situations they must be persuaded that their subjective judgments may be a decision maker's most valuable source of information.)

There are several procedural recommendations that may be helpful to designers and administrators of future Delphi exercises.

1. Interpersonal techniques, such as interviews and seminars, should be interspersed with the rounds of questionnaires and information feedback.
2. The source of a suggested item should be identified (for example, panel member number and basic biographical information), taking care not to compromise the anonymity of specific inputs.
3. Standardized scaled measures should be available to a respondent so that he can qualify his response to specific questions. Such measures are relative competence in a technical area, familiarity with a geographical region, or confidence in an estimate.
4. If a multidisciplinary approach is desired, respondents should be encouraged to consider all items but to make estimates only on those scaled descriptive phraseswith which he feels comfortable. For example, in these exercises it was helpful when respondents indicated their familiarity with a specialized area or the importance of an item even though they did not make probability estimates.

5. The panelists should decide through their suggestions and evaluations what items should be considered. The criteria for retaining an item for further evaluation should be made clear at the outset of the exercise.
6. A definite date on which the questionnaires are to be completed should be specified.
7. Personal comments and arguments submitted by respondents should be part of the information feedback.

The Sea Grant exercises have provided empirical evidence that the Delphi method can be useful in a broad range of situations. In addition we have the recommendations of exceptionally qualified panelists about potential applications of the technique. Perhaps most significant were the evaluations of the method's effectiveness in specific functions. These evaluations clearly support Turoff's observation that the primary objective of the Delphi process is to establish a meaningful communication structure and that the real issue is what communication process or combination of processes will be most effective, in terms of resources available, in examining the problem.[2/]

2/ Murray Turoff, "Delphi and Its Potential Impact on Information Systems," Paper No. 81 (Washington, D.C., Nov. 1971).

APPENDIX I

This appendix contains the guide that both researchers and decision makers on the broad panels were asked to use in making personal estimates of probability and an example of one of the evaluation questionnaires.

Guide for Personal Estimates of Probability

The results of the questionnaire showed a widespread familiarity with betting odds. A betting situation will therefore be used to develop the time estimates associated with the technical and social developments.

The systematic method for arriving at the timing estimates for each technical or social development is illustrated in one of the developments identified by the technical panel: "An information system for monitoring water quality, based primarily on airborne or satellite remote sensors, becomes operational for a region in the Great Lakes Basin."

You are asked first to select the year that divides the future into two periods in such a way that the development is just as likely to occur in either period. Visualize a movable pointer below a sequence of numbers representing years, as in Figure 1. Move the pointer until you reach a point where you would consider it a toss-up whether the development would take place before that year or after.

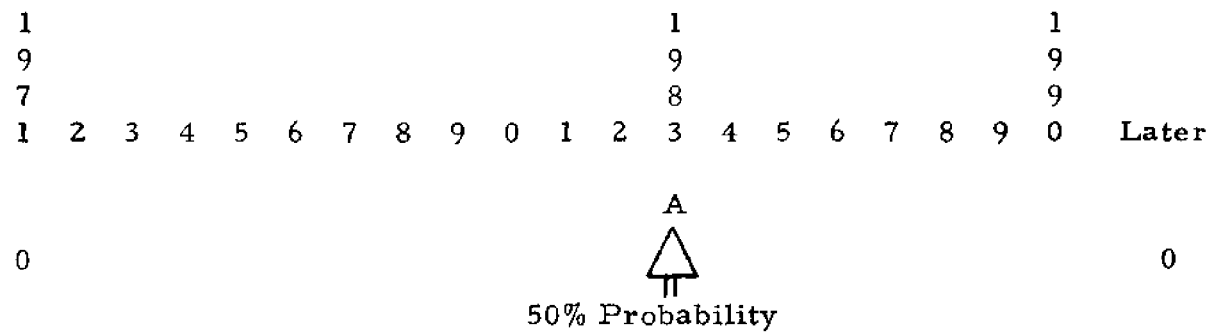

Fig. 1

In the example the pointer indicates that 1983 is the year that should be entered in column A of the Technological and Social Developments form, and it would be your 50 per cent probability date. It can also be described as the "1-to-1 odds" or "even chance" date. (If the pointer came to rest beyond 1990, you would record "Later" in column A and go on to consider the next development.)

Now consider the interval between 1971 and 1983 and move the pointer to a position where it is just as likely that the development would occur in the earlier period as in the later--with the stipulation that if the development were to occur after 1983 all bets are off. In Figure 2, the pointer rests at 1977, which is the year you should enter in column B and is your 25 per cent probability date. It can also be described as the "1-to-3 odds" or "rather unlikely" date.

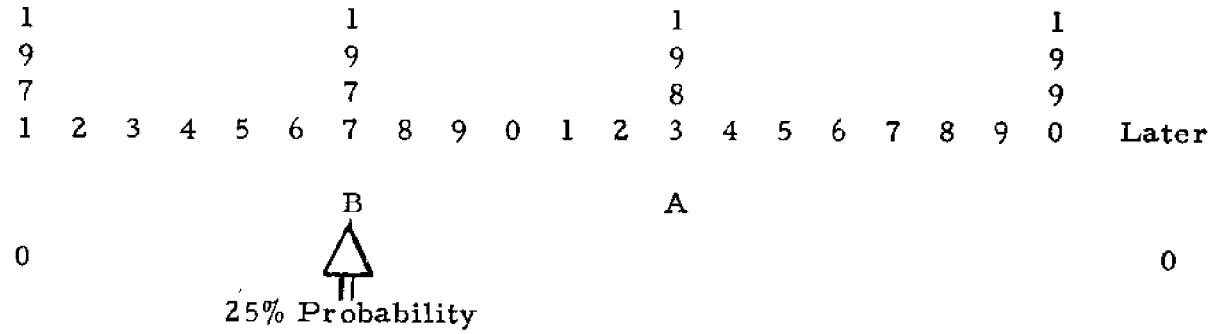

Fig. 2

Now consider the period after 1983 and again move the pointer to the point where you think you would select either side if you had to bet on the period when the development would take place. The stipulation here also is that if the development were to occur before 1983 all bets are off. In Figure 3, the pointer rests at 1986, and that is the year you should enter in column C and your 75 per cent probability date. It can also be described as the "3-to-1 odds" or "good chance" date.

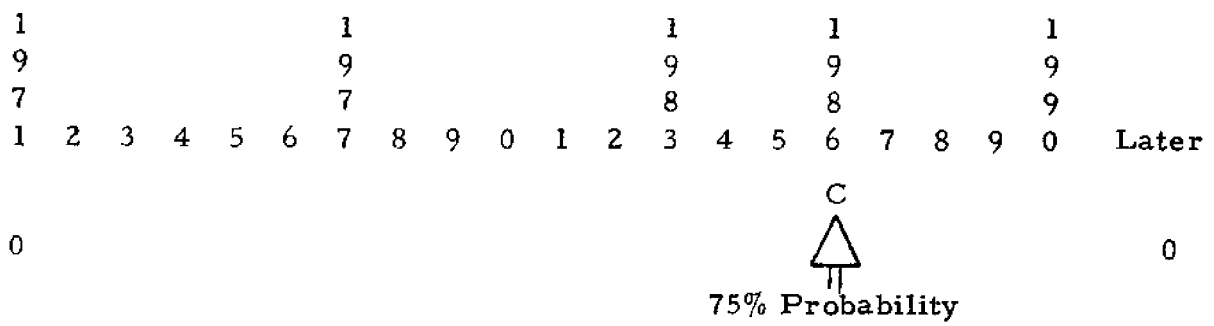

Fig. 3

For the moment lay aside the procedures and assessments which you have just made and approach the questions associated with Figure 4 with an open mind. The future is divided into four intervals as indicated by the reference lines on line X.

You are now asked to bet a significant amount of money on the interval in which you think the development will occur. You are given the first chance to select an interval, leaving the other choices to three other people. Is there any interval you prefer?

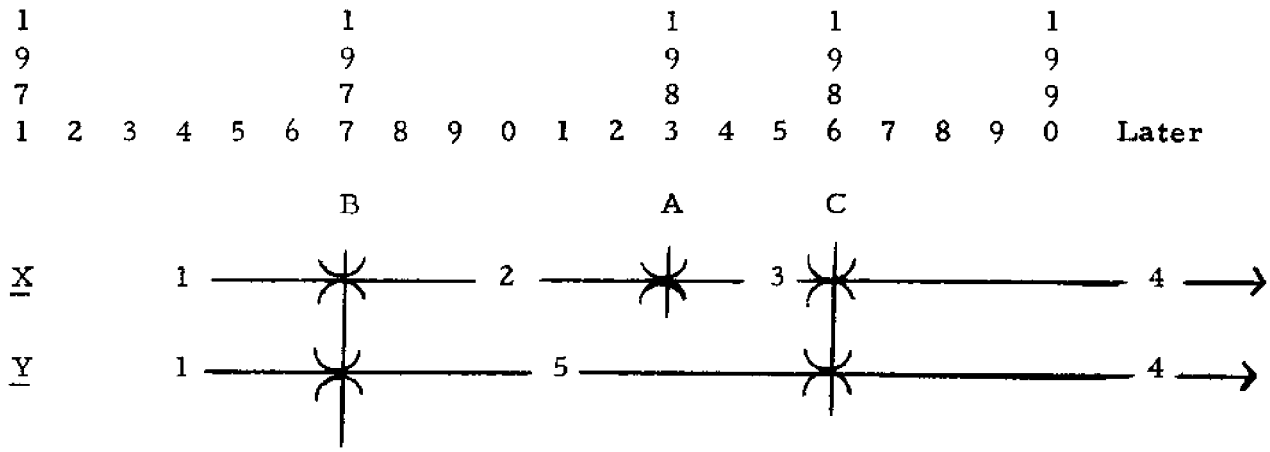

Fig. 4

Now consider the intervals indicated on line Y. One interval (5) extends from 1977 to 1986, spanning intervals 2 and 3 on line X. You are asked to bet a significant amount of money on whether you believe the development will occur during this interval or sometime before or after (that is, either between 1971 and 1977 or after 1986). Note there are only two choices, and you are given the first chance to select. Is one choice more attractive than the other?

If you would be as willing to bet on one interval as another in either line X or line Y then you have perfectly mastered this approach to probability assessments. If you do have a preference with either line, you are asked to run through the procedure once more and adjust your point selections.

After you have followed this routine mechanically for one or two developments, you will be able to estimate probabilities directly--as mental iterations of the routine just described are accomplished almost instantaneously.

Evaluation of Methodology--Broad Panels

The utilization of the intuitive judgments of experts is one of the most challenging problems facing a decision maker. The Delphi techniques are designed to provide a systematic method for obtaining informed judgments that preserves and complements the desired characteristics of conventional group meetings while overcoming some of their associated behavioral and administrative difficulties.

The respondents in the Sea Grant Delphi exercises are especially well qualified to evaluate methodology. Your recommendations could significantly improve and refine the Delphi method as well as widen its applications. A recent Wall Street Journal article which is critical of a version of the Delphi method, and an article on a "Conference Delphi" are included for your information.

Please mark the one phrase that comes closest to expressing your familiarity with the Delphi method prior to the Sea Grant exercises.

(1) ☐ Unfamiliar with the method

(2) ☐ Familiar with the method, but had not actively participated in a Delphi exercise

(3) ☐ Familiar with the method and had participated in at least one Delphi exercise

I read the "Guide for Personal Estimates of Probability" and used the suggested technique in making probability estimates for at least one development.

(1) ☐ Yes

(2) ☐ No

Rate each of the items below from two points of reference: This is your Experience in these exercises--designated (E); or this is your visualization of Potential applications of the Delphi method incorporating your suggested modifications and administered by a professional team of researchers and administrators--designated (P).

The rating scale:

-3,	+3	Strongly disagree, strongly agree
-2,	+2	Disagree, agree
-1,	+1	Somewhat disagree, somewhat agree

Circle or otherwise indicate your choice

		Disagree			Agree		
The Delphi techniques are effective in encouraging greater involvement by University personnel in the Sea Grant Program.	(E)	-3	-2	-1	+1	+2	+3
	(P)	-3	-2	-1	+1	+2	+3
The Delphi techniques are effective in obtaining, combining, and displaying the opinion of informed people so that their judgments can be utilized by decision makers and planners.	(E)	-3	-2	-1	+1	+2	+3
	(P)	-3	-2	-1	+1	+2	+3
The participation of people from the Grand Traverse Bay area on Delphi panels with University experts and researchers is an effective method of communicating information to regional planners and decision makers.	(E)	-3	-2	-1	+1	+2	+3
	(P)	-3	-2	-1	+1	+2	+3
If I were responsible for decisions involving regional planning and development, I would consider the use of a Delphi method.	(E)	-3	-2	-1	+1	+2	+3
	(P)	-3	-2	-1	+1	+2	+3
The use of a betting rationale--suggested in "Guide for Personal Estimates of Probability"--is helpful in making subjective probability estimates.	(E)	-3	-2	-1	+1	+2	+3
	(P)	-3	-2	-1	+1	+2	+3

The following information regarding positive and negative aspects of the Delphi techniques was obtained by collating the suggestions made by the technical panel. The source of each suggestion is indicated. You are asked to indicate your agreement or disagreement with each item, again using two points of reference--your experience in these exercises and your visualization of potential applications of the Delphi method incorporating your suggested modifications and administered by a professional team. The rating scale remains the same.

Positive Aspects

Source	Item		Disagree			Agree		
11	The advantages of a committee approach without the necessity of meetings	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
11	Easier to keep the whole picture in mind	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
12	Can get experts involved because preparatory work is done for them and they need to contribute only numerical estimates or brief statements	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
35	Gives one a broad perspective of potential technical developments related to a particular area	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
13	It brings to light ideas not likely to be considered by any one individual	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
17	Influence of strong personalities not as great as in other procedures	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
12	A systematic refinement of issues as exercise progresses	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3

Please suggest other positive aspects of the method.

_____

Negative Aspects

Source	Item		Disagree			Agree		
25	Not appropriate for complex, interdisciplinary problems	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
11	Time consuming	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
11	Tends to be tedious	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
11	No opportunity to change each other's opinion on the spot	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
11	No opportunity to hear reasons for various opinions	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
28	It lacks any relation to the economic realities of the times	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
28	One's judgment is frequently biased by constraints of the time	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
29	Very uninformed judgments are given to many questions	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
29	Answers usually lack cross-fertilization by the ideas of others, including some who are better informed	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
7	Can result in a manipulated and arbitrary consensus	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
13	Numbers provided by relatively uninformed people mean nothing	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
17	Chance for ambiguity in interpretation of some terms and questions	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3
18	It should provide an "overview" but it seems rather like a guessing game	(E)	-3	-2	-1	+1	+2	+3
		(P)	-3	-2	-1	+1	+2	+3

Please suggest other negative aspects of the method.

The following information regarding other applications of the Delphi techniques was obtained by collating the suggestions made by the technical panel. You are asked to indicate your agreement or disagreement with each item. The rating scale remains the same.

Suggested Applications for the Delphi Method

Source	Item		Disagree			Agree		
Administrator	Develop criteria for recruiting industries in a region--considering their desirability, feasibility, and impact on water resources	(P)	-3	-2	-1	+1	+2	+3
11	Budgeting an interdisciplinary program wherein individuals responsible for projects and specific areas consider the total budget	(P)	-3	-2	-1	+1	+2	+3
31	Education of practicing politicians by inclusion in Delphi exercises	(P)	-3	-2	-1	+1	+2	+3
31	Situations dealing with: (a) the future, (b) uncertainty, and (c) conflicting views	(P)	-3	-2	-1	+1	+2	+3
31	Long-range planning by a university--especially the nature of future growth	(P)	-3	-2	-1	+1	+2	+3
11, 12	Establishing program priorities within a program committee	(P)	-3	-2	-1	+1	+2	+3

Please suggest other applications.

APPENDIX II

This appendix contains a sample of comments made by panelists during interviews or in their responses to a Delphi interrogation. They represent a wide range of opinions concerning the techniques and may be helpful in the design of future Delphi exercises.

I was intrigued with the opportunity to see how other people responded to the questions.

A number of items are included that I know little about, but it was helpful just to think about them.

You people are going about it the wrong way. You should emphasize what is right instead of what is wrong. Problems may indeed be opportunities but by listing them under problems you bias the whole exercise.

Instead of bringing in new ideas you tend to perpetuate the old ways of doing things.

There is a wide range of issues, and groups will end up on different sides pretty much on the basis of their own economic interests. A medical doctor who has all of the patients he can handle is not receptive to programs that will have an adverse effect on his living environment. The quality of life was his motivation to migrate here and he is interested in preserving it.

Regarding methodology: I am not accustomed to thinking in terms of the probabilities of specific things happening in the future--I find it difficult to do.

I did not feel that I made any contribution in the exercise beyond what an interested individual in the Grand Traverse Bay region could have made--it did not take advantage of any specialized knowledge I might have. I became more interested in such things as recreational activities.

Mine is a very specialized area--I don't believe I have much to contribute.

Is there any implication for future research?

I was really interested to see fluorescence excitation by laser associated with an information system for monitoring water quality from an airborne platform. Several years ago I was working on this type of phenomenon in regard to a military application. It does seem that it might have application in this role.

If this does not produce participation and involvement I don't know what method will.

Relative to citizen involvement: There is a discrepancy between apparent participation versus actual involvement.

Use Delphi to identify new political arrangements to cope with environmental deterioration which threatens social economic structure, e.g., regional governments and alternative institutional arrangements.

I am interested in how scientific and technical knowledge is communicated to planners and policy makers.

Through their estimates and judgments can respondents be categorized as belonging to a particular social group, economic group, or political group?

Sir: Enclosed you will find our response to Round I of your project.

I realize that we are long past the deadline as set by you, however, it seems impossible for me to meet that deadline. Time, as little as I have, consumed in reading the material and the follow-up preparation eats up the hours in an amazing fashion.

If you wish me to continue on the panel on this basis I will do my best.

It takes hours to do conscientiously and is given no greater weight than the individual who treats it superficially.

City, county, and regional planning groups duplicate efforts which are costly. Communication is the big issue.

This technique does not give the decision maker the information that he really needs.

The exercise could have benefited from an orientation session covering the objectives and procedures and allowing an opportunity to ask questions.

I attend many conferences that include University experts and people in responsible positions in government and industry. I am reluctant to actively participate in the discussions until I have an opportunity to become familiar with the participants and the issues--which normally takes about a week. I am impressed with the potential of the Delphi method in providing a mechanism for introducing my suggestions at an earlier and more timely point in the conference.

Sometimes points are raised that I would like to think about prior to giving my opinion or would like to refer to supporting data that I know is available. In a conference, that usually means my opinion is not offered or considered--a deficiency which a Delphi methodology may improve. Can you tie an argument more specifically to its source?

I believe it is much more appropriate than gaming techniques, it involves a greater range of people, and it provides time to research and revaluate a position.

BIBLIOGRAPHY

Books

Brech, Ronald. Britain 1984: Unilever's Forecast--An Experiment in Economic History of the Future. London: Darton, Longman, and Todd, Ltd., 1963.

Bright, James R., ed. Technological Forecasting for Industry and Government. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1968.

Drucker, Peter F. The Age of Discontinuity. New York: Harper & Row, Publishers, 1969.

Jantsch, Erich. Perspectives of Planning. Paris: Organisation for Economic Cooperation and Development, 1969.

_____. Technological Forecasting in Perspective. Paris: Organisation for Economic Cooperation and Development, 1967.

Kahn, H., and Wiener, A.J. The Year 2000: A Framework for Speculation on the Next 33 Years. London: Collier-Macmillan, Ltd., 1967.

Articles and Reports

Bright, James R. "Evaluating Signals of Technological Change." Harvard Business Review, XLVIII (Jan.-Feb. 1970), 62-70.

_____. "An Academician's Introduction to Technological Forecasting," and "Some Insights from the Analysis of Past Forecasts." Papers presented at a conference entitled, "Technological Forecasting: An Academic Inquiry," Austin, Texas, Apr. 22-26, 1969.

Carter, Anne P. "The Economics of Technological Change." Reprinted from Scientific American, CCXIV (Apr. 1966), 25-31.

Cetron, Marvin J., and Dick, Donald N. "Technological Forecasting--Practical Problems and Pitfalls." Institute for Electrical and Electronic Engineers Transactions on Engineering Management, XVI (Nov. 1969), 161-72.

Chicago Tribune. "Traverse City's 'Shaggy' Waters Endanger Tourism." Oct. 26, 1970.

"Development Strategies--Upper Great Lakes Region." Annual Report of the Upper Great Lakes Regional Commission. Washington, D.C.: Government Printing Office, 1969.

Helmer, Olaf, and Rescher, Nicholas. "On the Epistemology of the Inexact Sciences." Management Science, VI (Oct. 1959).

Martino, Joseph P. "An Experiment with the Delphi Procedure for Long-Range Forecasting." Institute for Electrical and Electronic Engineers, Feb. 1967.

North, Harper Q. "A Probe of TRW's Future." TRW Corporation, July 1966.

North, Harper Q., and Pyke, Donald L. "Technological Forecasting in Planning for Company Growth." Institute for Electrical and Electronic Engineers Spectrum, VI (Jan. 1969), 30-36.

_____. "Technological Forecasting to Aid Research and Development Planning." Research Management, XII (July 1969).

Pyke, Donald L. "A Practical Approach to Delphi, Technological Forecasting, and Long-Range Planning." American Institute of Chemical Engineers, XV (Nov. 1969).

Sweet, David C.; Griffin, John M.; and Maggied, Hal S. Industries Suited for the Upper Great Lakes Region. Columbus, Ohio: Battelle Memorial Institute, 1970.

Turoff, Murray. "The Design of a Policy Delphi." Technological Forecasting and Social Change, II (Nov. 1970), 149-71.

_____. "The Delphi Conference." The Futurist, Apr. 1971.

_____. "Delphi and its Potential Impact on Information Systems." Paper 81. Washington, D.C.: Office of Emergency Planning Nov. 1971.

University of Michigan. Further Participation in the Sea Grant Institutional Support Program, 1970-71. A proposal to the National Science Foundation, Mar. 1970.

_____. Continued Participation in the Sea Grant Institutional Support Program, 1971-72. A proposal to the National Oceanic and Atmospheric Administration, II (Jan. 1971).

Umpleby, S. The Delphi Exploration: A Computer-Based System for Obtaining Subjective Judgments on Alternative Futures. Report F-1. Computer-Based Education Research Laboratory, University of Illinois, 1969.

Welty, Gordon. "A Critique of Some Long-Range Forecasting Developments." Proceedings of 38th Session of the International Statistics Institute. Aug. 1971.

Winkler, Robert L. "The Quantification of Judgment: Some Methodological Suggestions." Journal of the American Statistical Association, LXII (Dec. 1967).

_____. "The Consensus of Subjective Probability Distributions." Management Science, XV (Oct. 1968), 61-75.

_____. "The Assessment of Prior Distribution in Bayesian Analysis." Journal of the American Statistical Association, LXII (Sept. 1967).

Series of reports issued by the Institute for the Future presenting the findings of studies on the development of long-range forecasts of technological and social events

de Brigard, Raoul, and Helmer, Olaf. "Some Potential Societal Developments: 1970-2000." Report R-7 (to be published).

Enzer, Selwyn. A Case Study Using Forecasting as a Decision-Making Aid. WP-2.1970.

Enzer, Selwyn, and de Brigard, Raoul. Issues and Opportunities in the State of Connecticut: 1970-2000. Report R-8. 1970.

Enzer, Selwyn; Gordon, Theodore J.; Rochberg, Richard; and Buchele, Robert. A Simulation Game for the Study of State Policies. Report R-9. 1969.

Gordon, Theodore J., and Ament, Robert H. Forecasts of Some Technological and Scientific Developments and Their Societal Consequences. Report R-6. 1969.

Helmer, Olaf; Gordon, Theodore J.; Enzer, Selwyn; de Brigard, Raoul; and Rochberg, Richard. Development of Long-Range Forecasting Methods for Connecticut: A Summary. Report R-5. 1969.

Rochberg, Richard; Gordon, Theodore J.; and Helmer, Olaf. The Use of Cross-Impact Matrices for Forecasting and Planning. Report R-10. 1969.

_____. Research on Cross-Impact Techniques with Applications to Selected Problems in Economics, Political Science, and Technology Assessment. Report R-12. 1970.

Publications of the RAND Corporation
Santa Monica, California

Dalkey, Norman C. The Delphi Method: An Experimental Study of Group Opinion. RM-5888-PR. June 1969.

Gordon, T. J., and Helmer, Olaf. Prospects of Technological Progress. P-3643. Aug. 1967.

_____. Systematic Use of Expert Opinions. P-3721. Nov. 1967.

_____. Report on a Long-Range Forecasting Study. P-2982. Sept. 1964.

Helmer, Olaf. Analysis of the Future: The Delphi Method. P-2558. Mar. 1967.

_____. Systematic Use of Expert Opinions. P-3721. Nov. 1967.

_____. Methodology of Societal Studies. P-3611. June 1967.

Working Papers Issued by the Sea
Grant Advisory Services

Ludlow, John D. "The Delphi Method: A Systems Approach to the Utilization of Experts in Technological and Environmental Forecasting." Working Paper No. 3. Ann Arbor, Mich.: Bureau of Business Research, Graduate School of Business Administration, University of Michigan, Mar. 1970. (Multilith.)

_____. "Sea Grant Delphi Exercises: Techniques for Utilizing Informed Judgments of a Multidisciplinary Team of Researchers." Sea Grant Technical Report No. 5. Ann Arbor, Mich.: Sea Grant Office, University of Michigan, 1971. (Multilith.)

_____, and Braden, Patricia L. "Socioeconomic Development in the Grand Traverse Bay Area." Sea Grant Technical Report No. 8. Ann Arbor, Mich.: Sea Grant Office, University of Michigan, 1971. (Multilith.)

Warner, Katherine P., and Borton, Thomas E. "Developing Alternative Management Policies." Unpublished report, University of Michigan Sea Grant Office, 1971.