THE DELPHI METHOD: A SYSTEMS APPROACH TO THE UTILIZATION OF EXPERTS IN TECHNOLOGICAL AND ENVIRONMENTAL FORECASTING

Working Paper No. 3

by

John D. Ludlow
Research Associate
University of Michigan

FOR DISCUSSION PURPOSES ONLY

None of this material is to be quoted or reproduced without the express permission of the Bureau of Business Research
BACKGROUND OF THIS PAPER

This paper is based on research done at the Bureau of Business Research, Graduate School of Business Administration, the University of Michigan. The research was also financed by the Bureau.
SUMMARY

The need to extend planning horizons has become increasingly urgent with the accelerating pace of technological and environmental changes. Thus forecasting is of major concern, especially since conventional methods of extrapolation rely on data bases which become increasingly unreliable and irrelevant as the period of anticipation lengthens and the requirement for intuitive judgments becomes more crucial.

The Delphi method is a technique for systematically eliciting expert opinion and combining and refining it into a group judgment. Its most distinctive features are anonymity of the respondents, controlled feedback through the use of sequential interrogations, and a statistical summary designed to assure that every member of the group is represented in the response. The procedure minimizes some of the administrative, logistical, and behavioral barriers associated with conventional group deliberations and permits the investigator to screen out unproductive communications as the group moves toward a consensus. If the method is properly conducted it will increase the probability that all kinds of knowledge which bear on a problem will be represented and that the information feedback will be judged on its own merits.

In Delphi's most authoritative applications carefully selected experts have been used to supply information in the preparation of
explicit forecasts. Delphi procedures have been modified and combined with other techniques, such as TRW's\textsuperscript{1} mapping and T. R. Gordon's cross-impact matrix, to facilitate long-range planning.

Delphi methods are consistent with the following trends which are developing in forecasting and planning: (1) a search for techniques to improve intuitive thinking, (2) a search for integrative methods, (3) systems analysis, and (4) feedback mechanisms in which forecasting influences long-range planning.

A systematic integration of informed judgments with conventional quantitative methods such as cost-effectiveness comparisons, input-output analysis, and Bayesian decision theory will improve the dialogue between the staff technicians and the corporate decision maker and, at the least, will make the influence of intuition more explicit.

The Delphi method has been generally evaluated in the context of forecasting. However, the methodology need not have an orientation to the future; it can also be extremely valuable in using expert judgment for analyzing the past so as to improve future policy making and near-term deliberations.

The Delphi method and indeed the whole area of systematizing the use of expert judgment is in a primitive stage of development, and some

\textsuperscript{1} The firm of Thompson-Ramo-Woolridge is more commonly referred to as TRW, and this practice will be followed throughout this paper.
facets of the method still need to be critically analyzed and augmented. A larger role for Delphi methodology is seen in the development of higher-level objectives and goals and the concomitant policies and technologies that must be emphasized. Delphi's greatest potential lies in marshaling heretofore untapped reservoirs of expertise and in providing a mechanism for bridging communication barriers. Its role is to supplement and discipline other techniques rather than to serve as an alternative or competitive method.
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>The Basic Delphi Method</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>Delphi and TRW's Advanced Technological Forecasting</td>
<td>11</td>
</tr>
<tr>
<td>III</td>
<td>Delphi and Cross-Impact Methods</td>
<td>16</td>
</tr>
<tr>
<td>IV</td>
<td>Delphi: the State of the Art</td>
<td>22</td>
</tr>
<tr>
<td>V</td>
<td>The Outlook for Delphi</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>BIBLIOGRAPHY</td>
<td>42</td>
</tr>
</tbody>
</table>
FIGURES

1. Methods of displaying Delphi's statistical summary. 9
2. Convergence of opinions. 10
3. Scoring of TRW's additional parameters and overall summary. 14
4. Cross-impact method. 19
I

The Basic Delphi Method

Need for expert opinion

The need for marshaling expert opinion to extend planning horizons has become increasingly urgent with the accelerating pace of technological and environmental changes. Beginning in the 1940s the introduction of a body of analytic methods grouped under the heading of operations research made it feasible to think of many possible futures with associated probabilities and to provide some indication of the type of technological effort required to alter those probabilities. The concurrent development of computers and the refinement of operations research methods made possible incredibly complex calculations, but partly because of this complexity decision makers accepted the results only with reservations about their applicability. There are, of course, many elements involved in projecting future environments which do not lend themselves to quantification. Even when a rigorous formal mathematical model is available, judgment is exercised about input assumptions and output interpretations.

The expert has at his disposal a large store of background knowledge and a refined sensitivity to its relevance. In assessing the probability of a future event he will rely, of course, on some specific and explicit evidence. But he will also draw upon a vast body of potentially relevant
background knowledge that may be vague in its extent and deficient in its articulation. To the extent that judgments are arrived at by reference to such background knowledge, they can be called intuitive.

There are inherent dangers in resorting to the intuitive judgments of a single expert. James R. Bright cautions that technological expertise alone is not an assurance of good technological forecasting. The expert usually has no discipline for forecasting, and his opinion is vulnerable to the usual flaws of most personal opinion—limited personal experience, bias, and incomplete analysis. He is well aware of the physical laws which govern and limit progress in his field, and he is probably familiar with the significant research being conducted in his specialty. He will, however, tend to view a problem in terms of the boundaries of his own expertise, even though his is only one of several disciplines bearing on the problem. Perhaps most important, he is likely to have an imperfect appreciation of the process by which technological capabilities become economic reality.

Some of the disadvantages of consulting a single expert can be overcome by consulting a properly chosen committee or group of experts. In the area of technological forecasting, for example, the bias associated with a single specialty can be reduced and the likelihood of overlooking


important complementary research can be minimized when all pertinent
technology is represented on the committee. The performance of an
expert should also be enhanced when he can interact with other experts
in the same or in related fields. Even when a group of experts can't
reach agreement, the concept of brainstorming--the hypothesis that
among a large number of ideas advanced by talented people there will
be at least a few good ones--makes an attractive argument for the com-
mittee approach.

The committee approach, however, does introduce some admin-
istrative, logistical, sociological, and psychological barriers. In treating
complex problems it is not feasible to assemble experts from all the
pertinent disciplines and in the numbers required on a continuing basis.
Even if the administrative and logistical problems could be surmounted,
there remain formidable psychological barriers such as bandwagon
effects, face saving, personal confrontations, reluctance to risk ridicule
or to compromise an established reputation, unwillingness to abandon
publicly expressed opinions, and so on.

The interdisciplinary and interfunctional thinking required in fore-
casting future environments by means of the committee approach involves
problems of communication. For example, it is difficult to establish a
common language for exchanges of information among many disciplines,
each with its own peculiar jargon. There is also the tendency of individuals
to tune out when a viewpoint contrary to their own is expressed--partic-
ularly by someone not in the fraternity.
The output of group discussions is likely to be a compromise weighted toward the opinions of those in the group who are most articulate, most prominent, highest in rank, or administratively in the best place. The Delphi method, however, seeks to preserve the advantages of the group or committee approach to collective knowledge or judgment while minimizing the disadvantages.

The basic Delphi technique

The basic Delphi technique evolved from a series of studies conducted by researchers of the RAND Corporation initially under the guidance of Dr. Olaf Helmer and Dr. Norman Dalkey. The technique has been widely discussed since it was first described in a professional journal in 1963 and in a series of RAND reports beginning in 1964. The object of the method is to obtain the most reliable consensus of opinion from a group of experts through anonymous responses to a series of formal questionnaires interspersed with controlled opinion feedback.

The initial research was sponsored by the United States Air Force and was accomplished in the early 1950s. This experiment was designed to apply the expert opinion of a group of hypothetical Soviet strategic

---

3/ RAND, an acronym for research and development, is a nonprofit organization which works under contract for the U.S. Air Force. It was instrumental in developing the systems analysis technique, the military cost-effectiveness system, and the Planning, Programming, Budgeting system for the Department of Defense. The RAND Corporation has been active in investigations of technology transfer and technological innovation, and it is becoming a leader in the development of special forecasting techniques in social technology.
planners to the selection of an optimal U.S. industrial target system and to the job of estimating the number of atomic bombs of a given yield required to reduce munitions output by a prescribed amount.\footnote{4}{Olaf Helmer and Norman Dalkey, "An Experimental Application of the Delphi Method to the Use of Experts," \textit{Management Science}, IX (Apr., 1963).}

The method received its largest boost in general interest when it was applied by Gordon and Helmer to a study of forecasting technological events.\footnote{5}{T. J. Gordon and Olaf Helmer, "Report on a Long-Range Forecasting Study," paper reproduced by the RAND Corporation as a courtesy to its staff, P-2982, Sept., 1964.} The experiment was a trend-predicting exercise covering a period which extended as far as fifty years into the future. A sequence of questionnaires was used to elicit individual experts' predictions about six broad developments deemed to be of world-wide importance, namely scientific breakthroughs, population growth, automation, space progress, probability and prevention of war, and future weapon systems. Of the 150 experts who were approached, 82 responded; of these, 35 were members of RAND and seven others were RAND consultants. The Delphi method subsequently underwent significant refinement and increased in application, particularly for technological forecasting in industry.

It was also used by a variety of organizations for exploring policy decisions and as one of many tools in the planning process. Administrative measures including the streamlining of questionnaires and the use of computers have been employed to sharpen the technique.
An example of the Delphi method

Before considering recent applications and modifications of the Delphi technique it may be helpful to proceed through a Delphi exercise in a manner similar to that which has been used at technology and management conferences. 6/

Under operational conditions each expert should be selected for his particular knowledge of the field in which projections are being made. A series of rounds is conducted in which the selected experts are interrogated and in turn respond by correspondence, remaining anonymous to all except the chief investigator.

**Round 1.** Each participant is sent a questionnaire in which he is asked to name significant events such as inventions and scientific breakthroughs which he believes will occur in his field within a given period—the next fifty years, for example. The investigators edit these predictions to eliminate duplicate events, combine similar events, and retain only discrete and relevant events.

**Round 2.** Each participant receives the edited list of events and is asked to estimate the date at which there is a 50 per cent probability of realization of the event. The responses are summarized to indicate the median estimate of the group and the spread of the opinions. In treating the variability in the responses the basic Delphi method divides

---

6/ Two specific conferences were the Annual Technology Conference, "Technological Forecasting for Industry," held at the Lake Placid Club, New York, May, 1967, and the conference entitled, "Technological Forecasting: An Academic Inquiry," held at the Lakeview Inn near Austin, Texas, Apr., 1969. Both conferences were under the direction of Dr. James R. Bright.
them into quartiles. The interquartile range—the interval containing the middle 50 per cent of the responses—is considered the current consensus. In each subsequent round information feedback in the form of a statistical summary of the previous round is provided.

Round 3. Letters to the participants present the statistical summary for the group and ask the participants to reconsider their previous estimates. If the expert's revised response falls outside the interquartile range of estimates, he is asked to state briefly his reasons for his extreme opinion. The following typical third-round entry will illustrate this phase of the technique.  

The following question was posed: "In what year will the percentage of electrical among all automobiles in use reach 50 per cent?" This is a poorly designed question for developing a consensus about the market share that electric vehicles will realize in future years. It invites responses of "never," as well as a wide range of responses from those who feel that the 50 per cent level will be reached only very gradually. Such answers complicate a statistical analysis. Moreover, the question requires rather sophisticated judgments about the useful life of significantly different types of vehicles, how this will affect total vehicles in use, and the very definition of an automobile. Thus more appropriate wording of the question might be: In the year 1988 electric vehicles will make up what percentage of new car sales? An even better basis

\[7\] From a Delphi session which was conducted by Dr. Olaf Helmer at the First Annual Technology and Management Conference, May, 1967.
for developing a scenario might be gained if the respondents were asked to estimate values for the phenomenon or parameter of interest for several future points in time.

The responses to the third-round questionnaires included the following among the arguments to support a minority or extreme position: In favor of an earlier date, "Developments in nuclear power cost-effectiveness and in energy storage suggest an earlier date." In favor of a later date: "Pollution will force improvements." A study of the extreme (minority) arguments may indicate to the investigator that there is some ambiguity in the instructions, the question, or the event, and he may adjust these accordingly. Any reduction in ambiguity would tend to reduce the variability of subsequent responses.

The minority opinions are included in the questionnaires of subsequent rounds along with the consensus (statistical summary), allowing the experts to consider the arguments associated with divergent opinions in reconsidering their own estimates. The anonymity of the author ensures that the arguments will be considered on merit alone.

Experience has shown that as the rounds progress in most cases the consensus appears to narrow or to polarize around two positions. Figures 1 and 2 illustrate several examples of the statistical summary and the convergence toward a consensus.
Fig. 1. Methods of displaying Delphi's statistical summary.

Q - Q' = INTERQUARTILE RANGE  
M = MEDIAN
Fig. 2. Convergence of opinions.
II

Delphi and TRW's Advanced Technological Forecasting

In December, 1965, TRW adapted the Delphi method for use in its long-range forecasting of technology. The initial effort was primarily a test of the method's feasibility as an aid to more accurate trend extrapolation and for anticipating events not normally exposed by extrapolative techniques. The study involved a single panel of 27 of TRW's senior technical people and culminated in the development of a list of 400 technical events anticipated between 1966 and 1985 which were expected to have a significant impact on the company's products, services, and processes.

TRW was favorably impressed with the initial use of Delphi for technological forecasting and in the fall of 1967 launched a more comprehensive effort designated Probe II, which was built upon the experience in the first exercise and had the support and participation of key managers and senior executives. The 140 panelists were selected

from among the "most imaginative and creative members" of TRW's technical staff, and all divisions of the company were represented.

The panelists were provided with a set of detailed environmental assumptions upon which TRW bases its long-range financial plan. Fifteen specific areas were selected for investigation.

In addition to the forecast of specific events and dates, these four related parameters were developed using the Delphi method:

1. **An index of desirability** considered from the viewpoint of the consumer, i.e., the importance of the event to marketing

2. **An index of feasibility** reflecting both technical and economic considerations

3. **A timing refinement** in which the panelist assumes first that the events listed are certain to occur (probability = 1.0), and then he submits dates by which he thinks there is a reasonable chance the event will happen (probability = .1), a likely chance of occurrence (probability = .5), and the date by which the event is almost certain to occur (probability = .9)

---

9/ To develop the desirability and feasibility indexes individual panelists were given a choice of three descriptive words or phrases. For example, under desirability the choices were "needed desperately," "desirable," and "undesirable," and these were translated into a score of 1, 0, and -1 respectively. When the individual ratings are combined to form the composite evaluation by all panelists, a somewhat misleading impression of precision is presented, e.g., .86 for a given event.
4. A self-evaluation indicator based on familiarity with the technologies relevant to an event

Figure 3 illustrates the scoring of these additional parameters by individual respondent and gives an overall summary.

At the beginning of the exercise the panelists were asked to attend a briefing session at which general instructions were issued and background material describing environmental assumptions was distributed. From that point on the exercise proceeded in much the same manner as described for the basic Delphi technique. There was one major modification, however. Round 3 was limited to a resurvey of panelists who had rated their familiarity with a given event as excellent and whose judgment about the timing of the event lay outside arbitrarily established limits. There was reason to suspect that these experts based their judgment on information not available to other panel members, and the procedural innovation was an attempt to ferret out information not generally available to all panelists but which might have a significant bearing on one or more of the parameters related to an event. In addition to the normal statistical summary and the arguments of those holding an extreme position, the Round 3 questionnaire contained some specific remarks designed to challenge the expert about his position.

Although TRW's original exercise included a fourth round in which convergence toward a tighter consensus was sought, Probe II was terminated at the completion of Round 3 after a judgment that any gain to be realized from a fourth round was marginal. Probe II
### AN INDIVIDUAL RESPONSE

<table>
<thead>
<tr>
<th>Event</th>
<th>Familiarity</th>
<th>Customer Desirability</th>
<th>Producer Feasibility</th>
<th>Probable Timing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fair</td>
<td>Good</td>
<td>Excellent</td>
<td>Needed</td>
</tr>
<tr>
<td>E₁</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>E₂</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Eₙ</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

### SUMMARY

<table>
<thead>
<tr>
<th></th>
<th>Desirability</th>
<th>Feasibility</th>
<th>Timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>E₁</td>
<td>.25</td>
<td>-.29</td>
<td>.1----.5----.9</td>
</tr>
<tr>
<td>E₂</td>
<td>.67</td>
<td>.00</td>
<td>.1----.5----.9</td>
</tr>
<tr>
<td>Eₙ</td>
<td>.33</td>
<td>-.10</td>
<td>.1--.5--.9</td>
</tr>
</tbody>
</table>

Fig. 3. Scoring of TRW's additional parameters and overall summary.
resulted in 1,438 discrete events which will serve as "triangulation points upon which to base a map of TRW's future."

Results of the first exercise showed that the passive way in which the list of forecasted events was generated left gaps in some areas, and events forecasted to occur in the distant future influenced the near-term planners very little. In an attempt to remedy this situation a technique was developed which combined both normative and extrapolative methods of forecasting. It involves the development of a logic network which displays the sequence of technical developments that are not yet available but which must precede a forecasted event. It also involves a listing of all products or processes which may become feasible as a consequence of these technical developments.

TRW proposes to experiment with a similar mapping concept for the events developed in Probe II. This proposed experiment will ultimately be expanded to include:

A time-phased "shopping list" of all products or services likely to be of interest to TRW in the foreseeable future, all technological developments which are prerequisite to those products or services, all alternative products or services made possible by those technological developments, and, finally, significant anticipations in the political, economic, social, and cultural environments which may prove to be precursors of either the technological developments or their applications.

TRW's mapping concept requires that judgments resulting from evaluations made independently of one another be modified for mutual consistency in the light of interdependencies established during the mapping process.
III

Delphi and Cross-Impact Methods

T. J. Gordon, one of the authors of the Delphi technique, has come up with an innovation to answer the criticism that the Delphi method yields a set of linearly independent estimates of the future with the probability or date of each event estimated independently of the others. His research--labeled the Cross-Impact Method of Forecasting--is an attempt to develop a method by which the probabilities of an item in a forecasted set can be adjusted in view of judgments relating to potential interaction of the forecasted items. This, in essence, represents a quantitative approach to the formation of internally consistent scenarios.

Herman Kahn, one of the leaders in the field of scenario writing, has made the observation that scenarios are one of the most effective tools in lessening carry-over thinking. Scenarios are hypothetical sequences of events constructed for the purpose of focusing attention

10/ Mr. Gordon, together with Dr. Helmer, is currently at the Institute for the Future. He is a former Director of Space Stations and Planetary Systems for the Douglas Aircraft Company.

11/ Scenario writing denotes a technique which attempts to set up a logical sequence of events in order to show how a future state may evolve step by step starting from the existing situation. The technique has been pioneered by the RAND Corporation, the System Development Corporation, and the Hudson Institute and has been applied primarily to the exploration of potential military and diplomatic crises.
on causal processes and points of decision. In addition to treating the question of how some hypothetical situation might come about step by step, scenarios enhance the development of alternatives at each step for preventing, diverting, or facilitating the process.

In Gordon's cross-impact method Delphi techniques can be used to arrive at the events and dates that constitute the basic inputs to the development of scenarios. To the extent that the events and dates represent different disciplines and are arrived at independently, there is a strong probability that interaction between events and hence disciplines has not been adequately accounted for.

The interacting effects (among forecasted items) tend to be important not only because advances in one area are correlated with or spur advances in other areas, but also because various separate advances often allow for unexpected solutions to problems, or can be fitted together to make new wholes that are greater than the sum of their parts, or lead to other unexpected innovations.

To explain the cross-impact method fully it would be necessary to devise an example in some depth using quantities and mathematical formulas. It is hoped that some appreciation of its potential may be gained from the following brief description.

---


13/ Ibid.

Suppose that a set of developments is forecast to occur in the future with varying levels of probability. If there is a change in the probabilities associated with all other events when a specific event occurs we can speak of a cross-impact effect.

Operationally the events of interest, $E_1, E_2, \ldots E_n$, are arrayed vertically and horizontally in a matrix as indicated in Figure 4. The interaction or linkage between developments is expressed as a function or mode, as strength, and as diffusion time. The procedure recognizes three modes: enhancing, inhibiting, and unrelated.

The **enhancing linkage or mode** is that mode in which the probability of the second event is improved by the occurrence of the first event. This can happen in either of two ways: through an "enabling" reaction in which the first event makes the second seem more feasible or attractive, or through a "provoking" reaction in which the first event necessitates an effort to bring about the occurrence of the second.

The **inhibiting mode** is the reverse of the mode above. In it the occurrence of the second event is hindered by the occurrence of the first, and again this can happen in two ways. In a "disabling" reaction the occurrence of the first event makes a second event unfeasible or impractical; in a "deterring" reaction the first event necessitates effort to prevent the second.

The thrust of these modes is represented by plus, minus, and 0 respectively in the matrix. Thus in Figure 4 if $E_2$, "the feasibility of limited weather control," were to be realized, the impact on $E_1$,
FORECAST EVENTS AND PROBABILITIES FOR A GIVEN TIME PERIOD

<table>
<thead>
<tr>
<th>Events (E_i)</th>
<th>Initial Probability (P_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(E_1) One-month reliable weather forecasts</td>
<td>.4</td>
</tr>
<tr>
<td>(E_2) Feasibility of limited weather control</td>
<td>.2</td>
</tr>
<tr>
<td>(E_3) General biochemical immunization</td>
<td>.5</td>
</tr>
<tr>
<td>(E_4) Crop damage from adverse weather eliminated</td>
<td>.5</td>
</tr>
</tbody>
</table>

LINKAGE BETWEEN EVENTS
IN TERMS OF EFFECT ON PROBABILITY

<table>
<thead>
<tr>
<th>If this event were to occur</th>
<th>Then the probability of these events is</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E_1</td>
</tr>
<tr>
<td>E_1 (one-month forecasts)</td>
<td>0</td>
</tr>
<tr>
<td>E_2 (weather control)</td>
<td>+</td>
</tr>
<tr>
<td>E_3 (biochemical immunization)</td>
<td>0</td>
</tr>
<tr>
<td>E_4 (crop damage)</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Event E_i</th>
<th>Initial</th>
<th>Final</th>
<th>Shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>E_1</td>
<td>.4</td>
<td>.6</td>
<td>+ .2</td>
</tr>
<tr>
<td>E_2</td>
<td>.2</td>
<td>.3</td>
<td>+ .1</td>
</tr>
<tr>
<td>E_3</td>
<td>.5</td>
<td>.5</td>
<td>0</td>
</tr>
<tr>
<td>E_4</td>
<td>.5</td>
<td>.4</td>
<td>- .1</td>
</tr>
</tbody>
</table>

"one month reliable weather forecasts," and \( E_4 \), "elimination of crop damage from adverse weather," would be to increase the probability of their occurrence, and this is indicated by a plus sign. Event \( E_3 \), "general biochemical immunization," would be unaffected, the 0 reflecting an unrelated mode.

These modes presented in the form of matrices serve as a visual reference for a general understanding of the links or interaction between events; however, the notions of strength of linkage and time effect must be introduced into the analysis.

**Strength of linkage** is the relative effect of the occurrence of one event on the probability of occurrence of another event.

**Time effect** is the length of time it takes for the probability of an event to change because of the occurrence of another event. For example, if two events are strongly linked in the enhancing mode the time required for a second event to attain the higher probability, given the occurrence of the first event, can vary considerably.

The matrix is programmed into a computer along with the mode, strength, and time constant for the events paired by the analyst. The computer selects an event and, using random numbers, decides whether the event occurs. If the event occurs the remaining events are adjusted, and the process is repeated until all events are accounted for. This single run-through is repeated many times, and the events are ranked according to initial probability, final probability, and probability shift as in Figure 4. The indication by probability shift is, in essence, a measure of the suspected interactions.
Conceptually this iterative approach to the formation of internally consistent scenarios is similar to the iteration-through-synopsis approach used by Ronald Brech, in which he produced independent scenarios describing prospective developments in six different fields (demography, psychology, sociology, technology, politics, and the economy) and subsequently combined and modified them toward compatibility through an iterative process. Erich Jantsch makes the observation that the thought processes of an individual trying to arrive at a decision are often described as the intuitive development of separate scenarios with subsequent iteration through synopsis.

In Gordon's cross-impact method the ranking by probability shift indicates which items are most or least influenced by the external events represented by the remainder of the list. A possible application of this approach is to test the effect of policy decisions on the probability of a set of events.

Gordon believes that the work reported on is only indicative of a methodology of cross-impacts. Some suggested refinements include:

1. The possible use of the Delphi technique to reach a consensus among experts about direction, strengths, and successor-predecessor linkages.


16/ Erich Jantsch, Technological Forecasting in Perspective (Paris: Organization for Economic Cooperation and Development, 1967). Dr. Jantsch has been serving as a consultant to the Director for Scientific Affairs of the OECD since 1962.
2. Prior to the construction of a matrix, a relevance-tree exercise could be conducted to ensure that the field under investigation has been reasonably well covered and that the items listed are fairly equal in weight.

3. Testing forms of mathematical relationships other than the assumed quadratic form between the initial probability and final probability of events and the assumed linear relationship between strength of linkage and time remaining.

IV

Delphi: the State of the Art

The basic Delphi technique for eliciting expert opinion and combining and refining it into a group judgment has several features that distinguish it from more traditional methods of arriving at a consensus of experts. Among its most distinctive features are these:

1. **Anonymity** effected through the use of questionnaires. This procedure reduces the influence of certain psychological factors and enhances the probability that divergent arguments, if they are included in the information feedback, will be considered on their merit.

2. **Controlled feedback** through the use of sequential individual interrogations or rounds. This permits the investigator to minimize duplication, ambiguity, and unproductive communication and to proceed in an orderly way toward a consensus.
3. A statistical summary of group response designed to assure that every member of the group is represented in the final response.

4. Minimization of committee activity and some of its concomitant administrative, logistical, sociological, and psychological barriers.

Dalkey, in reporting on the results of experiments conducted in the spring of 1968 at RAND to clarify the information processes occurring in the Delphi interaction, treats some of the intuitively beneficial modifications of the basic technique suggested in other Delphi exercises. His discussion of the experimental results could serve as a partial summary of the state of the art. 17/

The experiments used upper-class and graduate students from the University of California at Los Angeles as subjects and general information such as that found in almanacs as subject matter. The general-information question was considered an acceptable surrogate for eliciting opinion on future occurrences because "the subjects did not know the answer, they did have other relevant information that enabled them to make estimates, and the route from other relevant information to an estimate was neither immediate nor direct." 18/


18/ Ibid., p. 19.
The two basic issues being examined were (1) a comparison of face-to-face discussion with the controlled-feedback interaction, and (2) a thorough evaluation of controlled feedback as a technique for improving estimates.

In the first issue, the negative conclusion that discussion displays no advantage over statistical aggregation was confirmed; more often than not, face-to-face discussion tends to make group estimates less accurate. The result thus strengthened the case for the application of the Delphi technique in areas of partial information.

The statistical summary fed back in most of these experiments was the group median and the interquartile range developed from individual single-point estimates. In two experiments the subjects were asked for distributional responses rather than point estimates and individual distributions were summed to determine the mean or median of the composite distribution. The results were statistically insignificant by themselves. Theoretically, however, a more accurate measure of consensus could be obtained by summing individual distributions than by analyzing point estimates. The method used by TRW, which yielded individual probability dates for the estimated time of occurrence of .1, .5, and .9, appears to have merit. This not only gives a better indication of the underlying distributions but seems to avoid some of the confusion inherent in translating an implicit probability distribution into an explicit point estimation.

One experiment was devoted to the hypothesis that the task of estimating is a skill that can be learned. Although the subjects were
fairly naive in the kind of estimates involved in the exercise, the results indicated that "although a discernible learning effect exists for individual responses, this effect is dominated by the effects of feedback and aggregation into a group response." 19/ The idea of developing expertise in the Delphi technique as a means of enhancing the exercise remains an intriguing one, however.

Experiments were conducted to compare the effectiveness of other types of feedback besides a statistical summary of the answers from the previous round. The main interest is in the introduction of the opinions of panelists who hold extreme views—statistically determined. Although none of the variances in group response with and without feedback of reasons were significant, it was concluded that "the addition of formulating and feeding back reasons for previously expressed opinions did not increase the accuracy of initial estimates or produce greater improvement on iteration." 20/ However, most critiques of substantive applications of the method state quite emphatically that the feedback of extreme positions is a crucial element in stimulating the experts to take cognizance of areas they may have neglected through inadvertence or ignorance. Indeed, as Gordon and Helmer caution, an essential feature of the method is to invite a respondent who disagrees with the majority to state his reasons for such disagreement.

19/ Ibid., p. 55.
20/ Ibid., p. 59.
thereby giving all members of the panel an opportunity to accept or reject such reasons and to reevaluate their judgments on the basis of whatever merits they find in them. Thus, a "far-out" forecaster whose judgment may prevail in the long run will have his opinion rejected or discounted only if he fails to justify it to the panel of experts.

In testing the value of some self-evaluation techniques (such as that used by TRW in Probe II) researchers concluded that individual self-ratings do not seem to be sufficiently accurate to allow the selection of a more competent subgroup, but that the average of individual self-ratings appears to be a useful indicator of the accuracy of the group answer.

Dalkey cautions that the experimental subject matter (almanac-type questions) and the subjects (college students) do not precisely match the area of interest (complex problems of estimation by experts), and therefore the results should be interpreted carefully before being applied to substantive exercise. This would appear to be particularly true in these additional experimental findings:

1. Female subjects were less accurate than males in their responses, and they were more likely to change their answers.

2. Students majoring in the humanities made more accurate estimates than students majoring in one of the hard sciences.

3. The degree to which accuracy was affected by the amount of time given to answer was translated into a surprisingly low optional time for answering.

V

The Outlook for Delphi

Administrative innovations

Several administrative innovations have been introduced into the Delphi process that will facilitate communication and enhance the environment for the exchange and processing of information. These include improvements in questionnaires that lessen the time required for the panelists to understand what is required of them and to respond, improvement in the structuring of questions to avoid ambiguity and compound questions, practical limits on the number of questions, and techniques for presenting the statistical summary together with a fair representation of extreme arguments and other feedback information.

In any complex operational exercise electronic data processing is desirable, particularly for developing and displaying the statistical summary. The data should be generated and recorded so that a data bank is established which can be readily updated and made consistent with various areas of interest. Other refinements have been suggested to ensure that panelists are selected wisely and their anonymity preserved. Effort toward streamlining administrative procedures will continue to be useful, but determining Delphi's best applications and integration with other forecasting techniques is more urgently needed.
Forecasting trends

Technological forecasting has provided the best forum for the Delphi technique because planning horizons must be significantly extended in anticipating technological change, and the data bases for traditional methods of extrapolation are becoming increasingly unreliable and irrelevant. Any forecast that is to be taken seriously requires a large intuitive input by an expert whose judgment is highly regarded. But a consensus of experts is even better than the judgment of a single expert.

Discussing Delphi in the context of technological forecasting is not too restrictive if the definition of technology includes the behavioral sciences as well as the physical sciences, if the definition of forecasting includes conditional or probability statements about the future, and if the combination of technology and forecasting treats diffusion and impact as well as significant technical achievements. In truth, the distinction between technological forecasting, long-range economic planning, systems analysis, and cost-effectiveness becomes increasingly hard to make.

There have been a large number of special techniques developed for technological forecasting. Jantsch, in his world-wide study of technological forecasting for the Organization for Economic Cooperation and Development (OECD), identifies more than one hundred

---

22/ Jantsch, Technological Forecasting.
distinguishable versions of, elements of, and formal attitudes toward such techniques which he groups under twenty different approaches in four broad areas: intuitive thinking, exploratory forecasting, normative forecasting, and feedback--the latter being an interaction of exploratory and normative forecasting. The dichotomy between the normative and the exploratory is a useful one and appears to be widely accepted. 23/

Normative forecasting reflects thinking which is oriented to a mission. Goals, needs, objectives, or desires are identified, and the analyst works backward to determine what capabilities exist or could be developed to meet the needs.

The goal-oriented concepts rest on the thesis that the future environment will determine needs, and these needs will determine the technology that society brings into being. Therefore, if one could identify future needs and their relative importance, he could predict what technology will emerge. Presumably, he could then decide what research and development should be supported.

Exploratory forecasting is oriented to opportunity or capability.

---

As Bright says, "On the whole, these techniques are based on developing the historic time-series for technological parameters and extrapolating them or relating them to produce a forecast of the technological progress under study."

Some variations of this technique include:

1. Single-function parameters based on trend extrapolations of single attributes like speed or weight

2. Multifunction parameters based on trends of complex combinations of attributes like thrust to weight

3. Precursor events in which technological development is related to a necessary prior event or capability

4. Envelope curves constructed from a family of individual curves which suggest the course of future technological development

5. Analogies which enable us to hypothesize that one technology may develop in the manner and rate of another

A discussion highlighted by Jantsch of significant trends in technological forecasting may be an effective vehicle for speculating on future applications of the Delphi method.²⁴/ Jantsch notes the

following four important trends in the development of forecasting
techniques, which he associates with the gradual switch toward
function-oriented, integrated forecasting and planning: (1) the
search for techniques to improve intuitive thinking as exemplified
by the Delphi model, (2) the adaptation of systems analysis techniques
to forecasting, (3) the search for integrative techniques which com-
bine the forecasting of technological development with the determination
of the effect such developments will have on the market, on economic
or industrial sectors, on the nation, or on society in general, and
(4) the search for feedback schemes which will permit forecasting
techniques to be applied to the procedures of creative thinking of the
future, i.e., the anticipation and evaluation of alternative possible
futures and the corresponding reorientation of today's action in a
feedback cycle.

Intuitive thinking

The experts relied on for intuitive thinking should be carefully
selected and have at their disposal the pertinent arrays of historical
time series, with projections which use sophisticated methods for
extrapolation and the quantitative techniques of operations research
for analyzing them. The intuitive element is introduced when the
experts consider the validity and quality of the input data, the
applicability of theories or methods, the impact of innovation and
other influences that resist quantification, and the weighting and
combining of separate forecasts.
Systems analysis

The Delphi method can provide a systems approach to the extent that various disciplines such as economics, the behavioral sciences, and technical fields are represented on a panel and the members of the panel consider both long- and short-run implications. Alternatively, scenarios developed independently by specialists can be reconciled through a Delphi exercise, using either these same experts or experts who are generalists trained in evaluating and selecting among alternatives.

Integrated techniques--cost-effectiveness

The cost-effectiveness approach advanced by RAND and the Department of Defense under Robert McNamara is in essence an integration of cost-benefit analysis with systems analysis techniques. The precise computer output of the analysis, however, often belies the very subjective judgments that enter into the effectiveness portion of the ratio. 25/

25/ The author was the program director of the Advanced Manned Strategic Aircraft (AMSA) Program (now known as the B-1A Strategic Bomber Program) from March, 1965, to August, 1967. He had executive authority for the management of the program and was the management and technical representative at high-level Air Force and Department of Defense discussions of strategic bombers. (This included deliberations on the bomber version of the F-111.) During this period the AMSA program was the Air Force's top priority research and development program and was one of the programs designated by the Secretary of Defense for special management procedures. These procedures included periodic reviews by the Designated Systems Management Group (DSMG) in which participation was held to a minimum number of people--usually an air staff briefing officer, the commanding general of Systems Command, and the appropriate program director, together with the secretary of the Air Force and his staff.
Delphi techniques increase the probability that all capabilities that contribute toward effectiveness will be included in the analysis and, further, that arguments will be freely expressed, accurately represented, and judged on their merits. Without these techniques some noteworthy aberrations in the cost-effectiveness method have occurred, resulting in significantly suboptimal decisions. A Delphi methodology is not designed to eliminate the role of the decision maker and his assessment of the relative values of political and institutional considerations, but rather to ensure that significant information—including nonquantitative and nonmonetary information—plays a fair part in preparatory deliberations.

Input-output analysis

Dr. Anne Carter, in discussing the relations between technological forecasting and the work of economists specializing in input-output analysis, reinforces the idea of increased emphasis on the search for integrative techniques.

Input-output tables detail industrial transactions in a given year through the use of a matrix in which buying patterns between industries are shown vertically and product sales distribution to various markets is shown horizontally. The matrix provides a simplified model of the

26/ Dr. Carter is Director of Research, Harvard Economic Research Project. Her views on integrating technological projections with input-output projections were taken from "Technological Forecasting and Input-Output Analysis," paper presented at the Environmental Forecasting Conference, Austin, Texas, Jan. 18-23, 1970.
economy in tracing the flow of goods and services between industries. This method affords a first approximation of a systems approach to determining the economic feasibility and impact of government or corporate planning. The initial impetus for the input-output method was given by the government’s economic planning for defense programs, but the method has subsequently shown promise for use in corporation investment analysis, product planning, and market research.²⁷/

Input-output tables are based on industry relationships as they existed in the national economy in 1963. Therefore, current and future input-output structures must be developed, using extrapolative and other techniques that can take into account new products and product mixes and can, as Carter suggests, include environmental variables such as air and water pollutants per unit of output. Interindustry structures are significantly affected by technology, and obtaining forecasts of new technology and innovative techniques requires assembling in some meaningful way the estimates of technical experts. Estimating rates of diffusion of new technologies calls for a consensus of economic and marketing experts. This input of expert opinion strongly suggests the use of Delphi as an integrating agent.

²⁷/ The February 17, 1970, edition of the Wall Street Journal reports that 50 major U.S. corporations have taken up input-output analysis in the past couple of years. Union Carbide Corporation has spent over $50,000 annually in the past five years on a computerized input-output analysis model. Companies update and enlarge segments of the tables that concern them most.
Bayesian decision theory method

The Bayesian approach to statistical inference and decision combines subjective judgments about the states of nature (a priori probabilities) with sample evidence to provide the basis for decision making or for a predictive distribution. Acceptance of the validity of the Bayesian concept is far from universal, hindered by theoretical debates among objectivists (on the side of classical statistical inference) and subjectivists concerning the validity of subjective prior distributions and by the reluctance of decision makers to accept the posterior component of the prior-posterior analysis. The crucial consideration is that it is being used, though in modified forms in many cases, by corporate decision makers.\(^{28}\)

Refining the techniques for eliciting and combining the expert opinion that goes into the subjective prior probabilities appears to present the best opportunity for increased effectiveness of the procedures and more general acceptance of the method. In important complex problems it may be desirable to consult several experts representing diverse areas of specialization. The problem then is to combine several distributions (expert opinions) into a single distribution (a consensus) to be used as an input (prior) to the prior-posterior analysis.\(^{29}\)

It is important to appreciate that the subjective judgments of some of the experts may be based almost entirely on objective statistical


\(^{29}\) The ideas on combining subjective probability distributions are taken from Robert L. Winkler, "The Consensus of Subjective Probability Distributions," Management Science, XV (Oct., 1968), 61-75.
(mechanical) procedures and evidence. The Delphi method, which recognizes the behavioral aspects of eliciting expert judgments in a series of feedback and reassessment iterations, offers a means of reaching a credible consensus for the prior probabilities required in the Bayesian decision method.

Feedback mechanisms

Application of Delphi to feedback schemes in which the forecasting of alternative possible futures leads to some reorientation of planning suggests that Delphi should have a role in making value judgments. Dalkey indicates\(^{30}\) that the validity of Delphi procedures in value judgments—in the sense that respondents are willing to furnish lists of objectives or goals, to allocate weights, and to accept a statistical aggregation of weights supplied by a group—is much more obscure than in factual judgments and has not been experimentally tested. Basically, however, the procedures appear to be sound. For example, an extension of TRW's desirability index of consumer interest to a desirability index for social and national objectives appears to be feasible with some refinements. In any event, it is in the area of normative forecasting and the selection of goals, objectives, and missions that some advocates see Delphi's most productive application. From a behavioral standpoint there appear to be cogent incentives for the large corporation to make its goals compatible with national and social goals so that "members of

\(^{30}\) Dalkey, "The Delphi Method," p. 73.
the corporation can best achieve their own goals by directing their effort toward the goals of the organization. 31/

Before using Delphi to establish a consensus on goals and objectives it seems necessary to develop predictions about what kind of future is most likely, together with alternative futures and related probabilities. Gordon's cross-impact matrix techniques could be utilized in developing internally consistent scenarios associated with the various futures. These procedures would help to focus on the most sensitive variables and related technologies by measuring probability shifts based on "what if" questions asked of a Delphi panel of experts or alternatively through simulation procedures based on prior probabilities and established interactions between significant events.

Having developed a consensus on probabilities associated with alternative futures together with the sensitivity of these futures to key events and related technology, the panel would be better prepared for dealing with goals and objectives that had some basis in reality. More important, they could provide a facilitating input into long-range planning by indicating which technology and policy should be stressed if the desired goals are to be achieved.

Extending the concept of feedback schemes a bit further, a relatively stable panel of experts available to a Delphi methodology over a period of time would be sensitive to environmental signals that would reinforce

or challenge their original assessment of the future and the policies and technologies to be pursued. The original output of the Delphi exercise could be expeditiously updated. 32/

Areas of greatest potential

Areas in which Delphi has the potential for playing an important role are those requiring a systems approach which includes many related disciplines and a long-range point of view. Such objectives as improving the quality of life in urban areas, estimating the commercial feasibility of alternatives to the internal combustion engine for motor vehicles, or an integrated program in the management of regional development are likely candidates for Delphi.

Situations in which Delphi could be most effective are those in which some members of the panel are experts on certain aspects of a problem, such as a specific technology, but not on all aspects of the problem. It is important to the exercise that the respondents be vitally interested in the problem and be willing to think about it long enough to ensure stability of panel membership. Although a substantive Delphi exercise has never been conducted in a large university, such a setting would be ideal because many disciplines would be represented. The university environment would also be a logical place for an evaluation by the panel members of potential refinements and modifications of the Delphi techniques.

In dealing with complex problems, such as the management of marine resources in the Great Lakes' basin, the university not only has experts who are eminently qualified in specific technical areas—for example, specialists on remote sensing techniques for mapping used in the analysis of lake ice formation and break-up, water pollution, beach configuration, underwater features, and shoreline processes; it also has experts who are generalists whose competence spans economic, social, legal, and institutional matters. The Delphi method suggests ways of integrating the expertise of the specialists with that of generalists who are experienced in choosing among alternatives.

In the initial rounds the respondents with very specialized knowledge could be segregated by their specialties. The early rounds could be regarded as programmed brainstorming sessions in which a host of ideas are contributed. Only the good ideas will survive subsequent rounds. The consensus for a particular specialized area is then developed with minimum noise. At this point the panel could shift to an interdisciplinary integrated configuration.

The impressions of those who have participated in substantive forecasting exercises using Delphi techniques are that the experience is stimulating and broadening. It is a method that tends to narrow the information gap between the technologist and the social scientist. It allows and, with modification, forces the technical expert to consider developments in closely related technical areas as well as in social, political, ecological, and economic spheres.
The Delphi method and, indeed, the whole attempt to systematize
the use of expert judgment is in a primitive stage of development, and
most of its procedures need to be critically analyzed and augmented.
There is some concern over the present statistical summary which may
contain a systematic bias and may not adequately represent the underlying
distributions. The idea of weighting evidence—not only to recognize the
dichotomy between objective and subjective evidence, but to weight the
sources--needs to be further explored. (In the self-evaluation approach
used by TRW the weighting scheme was, in essence, 1 for those
who rated themselves as knowledgeable in a given area and 0 for
those whose familiarity was less than excellent.) A more dynamic
configuration than that offered by point estimates would be desired if
the method were used in sales forecasting—a configuration that included
enough information to establish trends and momentum.

A Delphi methodology need not be exclusively oriented to the future.
Deliberation by experts and a systematic integration of their judgments
could be extremely valuable in analyzing the past; for example, the
significant developments and pressures that led to an involvement in
Vietnam that no one expected or desired, with an eye toward modifying
future policy and identifying warning signals. Problems in the present
or near future in which intuitive judgment plays a dominant role are
epitomized in the corporate capital budgeting process. A Delphi method
could improve a systems approach; at least it could make the intuitive
component less arbitrary.
Delphi's greatest potential is in applying the heretofore untapped reservoirs of special expertise that exist in an institution like the University of Michigan to interdisciplinary problems and in providing a mechanism for bridging communication barriers.

A methodology, such as Delphi, that will make it easier to establish a proper representation of experts thinking about complex problems from a systems viewpoint and that will provide an environment in which views can be freely expressed and considered on their merit warrants further consideration by the academic community.
REFERENCES

Books


Bright, James R., ed. Technological Forecasting for Industry and Government. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1968. The following articles were particularly useful:

Ayres, Robert H. "Envelope Curve Forecasting," pp. 77-94.


Journal Articles

Journal Articles--continued


Papers Reproduced by the RAND Corporation as a Courtesy to Its Staff


Papers Reproduced by the RAND Corporation as a Courtesy to Its Staff--continued


Papers Presented at Forecasting Conferences


Miscellaneous


7. RECORD # = b1369945.
CALL # = WP5300 1970 no.3.
TITLE = THE DELPHI METHOD : A SYSTEMS APPROACH TO THE UTILIZATION OF EXPERTS IN TECHNOLOGICAL AND ENVIRONMENTAL FORECASTING / BY JOHN D. LUDLOW.
ALT AUTHOR = Ludlow, John D.
ALT AUTHOR = University of Michigan. Bureau of Business Research.