

THE DELPHI TECHNIQUE: A REVIEW OF
THE LITERATURE

by

ARTHUR JOSEPH PELLETIER

B. A., Kansas State University, 1970

42-6074

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

College of Education

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1972

Approved by:

Herbert E. Kaiser
Major Professor

LD
2668
RH
1972
P45
c.2
Docu-
ment

TABLE OF CONTENTS

Chapter	Page
1. INTRODUCTION	1
2. OPINION, PREDICTION AND EXPERTS	4
OPINION	4
PREDICTION	6
EXPERTS	9
SUMMARY	11
3. SIMULATIONS, MODELS AND GAMES	13
SIMULATIONS	13
MODELS	15
GAMES	16
SUMMARY	16
4. FACE-TO-FACE INTERACTION AND DELPHI	18
FACE-TO-FACE INTERACTION	18
DELPHI	20
SUMMARY	26
5. APPLICATIONS AND REFINEMENTS OF DELPHI	28
APPLICATIONS	28
Values and Evaluation	34
Quality of Life	37
The Future	40
REFINEMENTS	46

Chapter	Page
SUMMARY	52
6. DELPHI IN EDUCATION	55
PLANNING AND INNOVATION IN HIGHER EDUCATION	55
FUTURISTICS	67
SUMMARY	70
7. EVALUATION	71
STRUCTURED FACE-TO-FACE INTERACTION AND ANONYMOUS QUESTIONNAIRES	71
VALIDITY	72
RELIABILITY	72
CONSISTENCY	73
SUGGESTED IMPROVEMENTS FOR FURTHER RESEARCH	76
REFERENCES	78

Chapter 1

INTRODUCTION

Institutions of higher education, as do most organizations, seem to be somewhat dependent, at one time or another, on the results or decisions of committees, commissions, or panels of supposedly knowledgeable individuals. There seem to be committees for almost any type of activity, such as allocation of student funds, selection of a new chancellor, consideration of tenure applications, or the disposition of one's parking ticket. Surely, the reader can think of even more examples of this type of human interaction in colleges and universities. However, it is not the intention to discredit the concept of the committee, since they do seem to be necessary, especially when a large number of individuals will be affected by a single decision, or a series of decisions. There appear to be certain disadvantages, however, in the face-to-face interaction that occurs in committees; these have been disclosed in a large number of investigations by psychologists and sociologists interested in the structure, process and behavior of small groups. It may be effectively argued, therefore, that these disadvantages, which seem inherent in the nature of face-to-face interaction, deleteriously affect the eventual decisions in a number of different ways.

In recent years, researchers in a series of studies at the RAND Corporation in Santa Monica, California, have attempted to correct some of these disadvantages of in-person, face-to-face interaction. A technique was devised and later refined based on these studies, which eliminates face-to-face com-

mittee activity altogether, and replaces it with a carefully prepared program of sequential, anonymous, individual questionnaires, alternately interspersed with opinion feedback (derived by computing a consensus from the earlier segments of the program) and additional information. This technique was designed by Olaf Helmer and Norman Dalkey, and subsequently named (by philosopher Abraham Kaplan) the Delphi Technique (Dalkey, 14).

Since its introduction, Delphi has been applied in a variety of studies, of which a few have been somewhat diverse from its initial application. Among other things, the technique seems to have prompted, or at least contributed to, the recent, intense interest in the future that has invaded the sphere of education in the United States. Delphi is also being used in the business world, in teacher education, in defining institutional goals, and in the area of student affairs, to name just a few. In essence, then, Delphi may be utilized in any context where it seems appropriate to seek a consensus of opinions on a particular topic from a group of knowledgeable individuals (Helmer, 35).

A review of the literature on this intriguing and unusual technique will be presented by tracing its origins from as early as 1938 and 1950, to its development around 1953, through its evolution via a number of studies during the 1960's, to a new development only a few years old which attempts to eliminate some of the weaknesses in certain applications of the technique, called Cross-Impact Matrix (Gordon and Hayward, 24). Also to be discussed will be some of the various applications of Delphi, primarily in the general area of education (particularly higher education) as well as other areas. The technique, despite some shortcomings in particular applications, is shown to be a viable and effective decision-making alternative to the everyday committee, either by replacing it altogether, or by being used in conjunction with

it.

No attempt was made to describe in detail all of the experiments and studies that have been conducted utilizing the Delphi Technique; many are so complex in their nature, and so detailed in their discussion of theory and method, they almost defy synopsis. Further details on particular studies or experiments may be found in individual articles listed in the references. Most of these articles are, in themselves, a complete and thorough undertaking, and many include their own references. In the interest of smooth reading, few statistics, graphs or charts were presented except where they were considered to be necessary for description or clarification. Those experiments or studies cited in the text were given in the form of short descriptions of purpose and conclusions reached.

The initial discussion presented some of the terms and concepts encountered in Delphi literature, and was followed by a review of some of the other decision-making and forecasting techniques, of which Delphi is only one. The actual technique, itself, was discussed together with some of the advantages and disadvantages with respect to face-to-face interaction. Some of the experimental applications of Delphi were reviewed and the resultant refinements as suggested by these experiments were followed by a discussion concerning some of the applications of the technique to higher education and education in general. The discussion concluded with a general discussion of the technique, including a limited critique and suggestions for applications and improvements.

AL
7/24
24

Chapter 2

OPINION, PREDICTION AND EXPERTS

In the literature on Delphi, one encounters a number of terms and concepts which, though probably familiar, seem to acquire different connotations with respect to the Delphi Technique. In particular, three of these concepts will be discussed in the present chapter: 1) that type of information which lies between fact and speculation--opinion, 2) prediction, and 3) the use of individuals who are especially knowledgeable in their fields--experts.

OPINION

There appear to be three kinds of information that play a role in either individual or group decision-making. First, there are assertions that are highly confirmed with a great deal of evidence, which are usually classified as "knowledge". Second, there is that material which has little or no evidential support, which is usually called "speculation". Between these two extremes there exists a wide area of material for which there seems some basis for belief but which is not sufficiently confirmed to warrant its being considered knowledge. Dalkey (15) preferred to call this area of material "opinion", a basic characteristic of which, as opposed to more "solid knowledge", is evident in the fact that one is likely to receive a divergence of responses upon interrogation of several equally competent individuals. From the position of the decision-maker, this could, and at times does, create a problem of how

best to utilize these estimates or opinions in developing and formulating his policies (Dalkey, 11). As Dalkey (15) stated, "At all events, policy decisions will be made on the basis of opinion, and there remains the separable question whether techniques exist for refining it."

Helmer (35) agreed with Dalkey. There are many cases, he contended, where decisions have to be based on the intuitive judgment of the available experts in a particular field, either because no satisfactory theory has yet been developed, or because judgment must be sought as a matter of principle. The latter is due to the topic involving moral as well as factual aspects, and thus preferences as well as hard data.

"Once we have recognized that reliance on expert opinion at times is unavoidable, we need to give some thought to the question of how to obtain such opinions most efficiently and most reliably" (Helmer, 35). *

There seem to be a number of traditional approaches for the utilization of opinion, or what Dalkey (13) referred to as "opinion technology". The first approach is to select the best qualified expert and seek his opinion. The second approach is to attempt to include opinions, in addition to available knowledge, in a more comprehensive "conceptual framework". A third approach is to rely on the historical tautology of "n-heads are better than one".

The basis for the n-heads rule is that on a given question, there will be at least as much relevant information in n-heads as there will be in any one of them. Likewise, it is also a tautology that there could be as much misinformation in n-heads as there is in one. However, it is not tautological that a technique exists for extracting the information in n-heads and combining it to form a more reliable opinion. One must be cautious, though, for it could very well be the misinformation that is being combined, and into a less

reliable opinion (Dalkey, 15). This suggests, according to Dalkey (11) that there may be an optimal size of group for a given kind of estimation, and that part of the group estimation process should be concerned with the active suppression of such misinformation. The approach represented by this tautology is none other than the usual committees, commissions, panels or boards, which are utilized when dealing with opinion or when decisions must be made using partial information (Dalkey, 14).

In concluding this review of opinion, and its use in decision-making, a statement by Rescher (44) seems very appropriate: "The systematic (and preferably structured) utilization of expert opinion and speculation is perhaps the principal and most promising forecasting tool in the technological-scientific-social domain...."

PREDICTION

Prediction-making is considered a fundamental aspect of technological, military, commercial, political and social planning. Forecasts of relatively short-range events of, for example, the next twenty-four hours, or next year, or even trends of the next ten years are usually sufficiently accurate to be of demonstrably practical use. However, as the period of concern is extended into the distant future, uncertainties tend to multiply, the confidence in prediction declines and many planning decisions must, of necessity, rely on intuitive judgment (Gordon and Helmer, 23).

Decision-makers invariably are forced to rely upon predictive ability, and thus usually surround themselves with staffs of expert advisors, whose usual roles are that of "predictors", that is, those who are able to sketch out a general trend or direction of future developments (Helmer and Rescher, 26).

There were two studies, reported in 1938, that purported to study prediction within the context of social and technological events. First, McGregor (40) was primarily concerned with 1) the validity of predictions, 2) the certainty with which predictions were expressed, 3) the effect of ambiguous stimulus situations on the predictive judgment, and 4) the influence on prediction of such subjective factors as wishes, attitudes, information, and "idiosyncratic" personality factors such as pessimism, optimism, cautiousness and skepticism.

McGregor termed predictions as psychological inferences made in the future tense, and that these predictions are influenced by subjective factors which comprise the unique personal characteristics of the predictor.

It was found that the influence on prediction of these subjective factors depended upon the importance of the predicted event (intensity of attitudes and wishes) for the predictor, and that the prediction was further affected by the degree of ambiguity in the situation. As McGregor stated, "If either importance or ambiguity were zero, the influence of subjective factors...would presumably be zero...[and]...if both ambiguity and importance were maximal, subjective factors would be completely determinative of prediction."

It would appear that from the above statements, one could infer that in the case of long-range forecasting, for example (which would concern distant future events) where the ambiguity of the situation would presumably be high, and the importance of the events to the predictor relatively low, the influence of subjective factors would still have an effect on the prediction, but the amount of that influence would lie somewhere between the two extremes suggested by McGregor.

McGregor concluded that there is a low, yet significant, correlation between the amount of knowledge (or information) and lack of prejudice. How-

ever, the person with a quantity of information is not necessarily the most correctly informed; "it is the nature of one's information that is determinative, not the amount."

The study conducted by McGregor suggested some effects of several personality factors on predictions. He contended that the predictions of the confirmed optimist were in accord with his wishes, while those of the pessimist were consistently in opposition to them. In addition, the predictions made by a cautious individual would have a lower probability of occurrence than if they were made by a less cautious individual. Finally, the predictions of the skeptic were more likely to be influenced by objective conditions, whereas one who was somewhat less skeptic would more readily accept less objective discussions and opinions. Any one or all of these factors would tend to produce a prediction that would be less than objective.

Cantril (6), writing on the "Prediction of Social Events" in a later issue of the same journal, felt that "predictions help to give structure to present experience, providing the individual with the psychological closure he craves." *good*

Cantril found that people were more sure of their judgments regarding the direction of the outcome of an event than they were of the date of that outcome, and that those individuals whose attitudes were noted to favor a certain outcome for an event, tended to forecast the desired outcome.

Along the same lines indicated by Cantril, is the objection that some predictions (especially those solicited by the Delphi process) might be of the self-fulfilling or self-defeating variety. That is, the statement of a positive attitude toward the outcome of a particular event might strengthen the general attitude toward that event and thus fulfill the prediction. Likewise, a statement of potential negativism might serve to boost interest and thus re-

sult in the prediction becoming self-defeating. Aside from this, there is the possibility that a predictor's answer might be biased by his expectation that the announcement may have an effect on the truth of the prediction's content. In this case, the respondent would cease to be an objective predictor, and become a manipulator of the future. However, there seems to be no evidence of a logical circularity in this instance (Rescher, 43).

In concluding this discussion of prediction, a statement by Rochberg (44) seems to adequately assess the situation: "It is obvious that almost any attempt to study a social or psychological phenomenon will affect the phenomenon being studied...[and]...it is equally clear that almost any attempt to make public predictions concerning a future social or psychological situation will affect future events and hence possibly affect the event predicted."

EXPERTS

Traditional social science methods seem to be inadequate in effectively dealing with the increasing complexity of forecasting the consequences of alternative actions and, therefore, in furnishing a useful planning aid to decision-makers. However, to remedy this situation, new methods have been developed, one of which is the systematic utilization of expert opinions (Helmer, 30).

Expert judgment may be solicited in any planning operation in which selection among several alternative courses of action is necessary, due to the absence of a theory which would evaluate the consequences of the proposed courses of action and select a preferred alternative by traditional procedures (Brown, 4).

As Helmer (30) stated,

In view of the absence of a proper theoretical foundation

and the consequent inevitability of having, to some extent, to rely on intuitive expertise...we are faced with two options: We can either throw up our hands in despair and wait until we have an adequate theory enabling us to deal with socio-economic and political problems as confidently as we do with problems in physics and chemistry, or we can make the most of an admittedly unsatisfactory situation and try to obtain the relevant intuitive insights of experts and then use their judgments as systematically as possible.

In the utilization of experts, Helmer (30) suggested three rules that ought to be followed: 1) Use wisdom in selection of the experts, 2) create the proper conditions for maximum performance, 3) exercise caution in deriving a single combined position from various opinions when several experts are available.

Regarding the first rule, the selection of experts is usually based on their prior record of past performance, and use of their judgment does not seem incompatible with scientific objectivity. If their prior record is unknown, however, reliance upon their predictions may be objectively justified on the basis of general background knowledge as to their reputation as experts. Furthermore, the objective reliability of experts' predictions appear to exhibit considerable agreement with one another, which, if they are independently made, would seem to preclude their being considered subjective whims (Helmer and Rescher, 26).

Fortunately, when the expert is unable, due to the influence of chance factors, to make precise predictions, he can be expected to indicate the major contingencies on which future developments will hinge, and to provide personal probabilities with respect to these contingencies (Helmer and Rescher, 26). However, the expert's judgment must be evaluated, its likelihood of success appraised, and improvements attempted on it (Kaplan, et al., 36).

As Gordon and Helmer (23) stated, "...it would seem that any improvement...that could be achieved by replacing casual guess with the controlled

use of intuitive expertise would be desirable because of the benefits that... might...[be derived]...from it."

SUMMARY

In the present chapter, three concepts encountered in the literature on the Delphi Technique have been reviewed: opinion, prediction, and experts. A brief summary follows.

It was noted that opinion was found to refer to that "middle area" of information between speculation and knowledge, and that it usually has some basis for belief, but not in sufficient quantity to be considered knowledge. It was shown that decision-makers must, at times, rely on opinion either as a matter of principle or because the topic involves moral aspects, and therefore preferences. How best to obtain such opinions was reviewed, and it was noted that systematic combining of opinions, particularly from experts, was preferred, but that caution must be exercised in doing so.

In reference to prediction, it was shown that this was a topic of concern as early as 1938, and that the two studies undertaken at that time purported to investigate subjective factors related to prediction, as well as personality factors. It was found that ambiguity of the situation, importance of the event to the predictor, the nature of the information possessed by the predictor (not necessarily the amount) as well as the general outlook and attitude of the predictor were all variables to be considered in relying on predictions.

The discussion then centered on the use of experts and intuitive expertise. It was shown that in the absence of adequate theory, reliance must, at times, fall on the systematic use of experts and expertise, although it does not seem clear exactly what constitutes an expert. It was also noted

that the selection of experts should be handled wisely, and that if they are unable to give precise predictions, they can provide other useful information.

As can be seen, it is difficult to talk of these three concepts as if they were exclusive of each other, when in actuality, they are interrelated. One statement, all inclusive, might serve to demonstrate this characteristic. Decision-makers must at times utilize predictions, and when this is the case, the use of experts and their opinions seems justified but only when these opinions can be combined in a systematic way because there will be different responses from different experts on the same topic, and several heads are better than one.

The discussion will now turn to the general group of decision-making techniques to which Delphi belongs. Again, however, the attempt will be made to discuss three interrelated techniques as though they were exclusive of each other.

Chapter 3

SIMULATIONS, MODELS AND GAMES

The Delphi Technique is only one of a number of techniques that may be utilized by decision-makers. The present discussion will concern three other techniques that are encountered in the literature on Delphi: 1) the use of simulation exercises, 2) models, and 3) operational gaming. *There are more equally valid is, cross impts*

SIMULATIONS

In the preceding chapter, the use of expertise was discussed and it was seen that Helmer (30) suggested three rules concerning experts, and the first rule, concerning the selection of experts was considered. The second rule, that an expert should be placed in the right conditions in order to perform well, seems more appropriate for the present discussion. The second rule means that a first consideration in the utilization of expertise, would be the formulation of an appropriate model which would serve to clearly communicate the problem to the expert and then receive his answer without risk of misinterpretation (Helmer, 30). In other words, as Helmer (30) stated:

and In order to provide access to intuitive knowledge that may not yet have been recorded, an expert's performance would be enhanced most significantly by placing him in a situation where he could interact with other experts in the same field or in related fields covering other aspects of the same problem.

One way of encouraging such interaction would be to place the experts in a laboratory situation where they would be required to participate in a simulation exercise (Helmer, 30).

It seems that a kind of conceptual transference occurs in a simulation exercise; instead of directly describing a situation, each element is simulated by the substitution of a mathematical or physical object for the real one and simulative relations for those that actually exist. For example, a planning operation could be simulated by a set of make-believe decision-makers who, by playing roles in a laboratory "game", would go through the decision-making motions that their real-life counterparts would be expected to actually perform (Helmer, 30). Helmer (30) stated it another way:

In a simulation model, instead of formulating hypotheses and predictions directly about the real world, it is possible instead to formulate them with reference to the model. Any results obtained from an analysis of the model, to the extent that it accurately simulates reality, can later be translated back into corresponding statements about the real world.

Reported experience with simulation models seems to suggest that they can be instrumental in motivating participants to effectively communicate with one another. In this way, they would be learning more about the subject matter by viewing it through the eyes of persons with different backgrounds and skills, as well as acquiring an integrated overview of the problem topic. The stimulation resulting from collaboration on the employment of a simulation model is particularly powerful when simulation takes the form of an operational game where the participants act out decision-making roles. This exposure, within a simulated environment, to a conflict situation involving an intelligent opposition, compels the "player", no matter how narrow his specialty, to consider many aspects of the situation that would not have normally influenced his opinions to the same extent had he worked in isolation (Helmer, 30).

Helmer (33) stated that the need is great for interdisciplinary mathematical models, especially in the social-economic-political development areas, and felt that progress here has been retarded due to the reliance on uni-

disciplinary models. The discussion shall now turn to this topic.

MODELS

Writers on operations research seem to maintain that the one major element in dealing with a problem is to construct and use an operational model of the system, incorporating measurements of factors such as chance and risk, with which to predict and compare the outcomes of alternative decisions, strategies or controls (Quade, 41). As Quade stated:

The model must represent or simulate the essential features of the situation under study. Such a model may take many forms, but the most useful, certainly the most used, and often the only types considered, are mathematical, expressing through a set of equations or a computer program the effectiveness and costs of alternative actions as a function of the variables one has under control. By operating with the model, either analytically or numerically, the consequences of alternative choices are determined and a preference identified.

It seems particularly difficult to create such a tried, but restricted model when the predominate factors are political and social in nature. As an example, goals in the social sphere seem to be, at times, obscure and conflicting, and without a well-defined structure (Quade, 41).

The primary function of a model is to predict and compare; to provide a logical way to forecast the outcomes that tend to follow alternative courses of action, as well as to indicate a preference among those actions. Reliance on expert judgment and intuition seems to be crucial to every decision process. This reliance is evident in every aspect of operations, from deciding the more fruitful approach, to designing the model, to determining the facts, to interpreting the results. One of the virtues of model building is that it provides a systematic, explicit, and efficient way to focus the required judgment and intuition (Quade, 41).

By introducing a precise framework and terminology, a model may serve

as an effective means of communication, enabling various experts to exercise their judgment and intuition in a context that is well-defined and in proper relation to each other. In addition, a model provides feedback to guide the participants in the revision of their earlier judgments. These seem to be the essential features of a model in its role of supplying a path from hypotheses to prediction, not simply the explicitness with which it purports to represent the real world (Quade, 41).

GAMES

Just a step away from the traditional model and now an accepted operations research technique, is that method known as "operational gaming". These are exercises in which the participants interact by playing roles that simulate individuals, factions in a society, or sectors of an economy. The predictive quality of gaming, however, seems to be a function of the insight intuitively provided by the participants. By allowing for the introduction of judgment at each step, a game provides an opportunity to take into account intangibles often considered completely beyond analysis (Quade, 41).

However, Quade (41) contended that gaming retains the representative features of the traditional model, and suggested that a broad view be taken by accepting as a model, "any device that provides a logical means to predict and compare the outcomes of alternative actions."

SUMMARY

Three of the group of methodologies to which Delphi belongs were presented in this chapter, namely, simulations, models, and games. A brief summary follows.

It was seen that simulation exercises encourage interaction among ex-

perts, thus motivating effective communication, aiding in learning more about the topic by forcing them to consider other viewpoints, and in the form of operational games, requiring them to consider other facets of the problem that could be overlooked by working alone.

It was noted that the function of a model is to predict and compare alternatives, and that it is very useful in providing an efficient and systematic means of focusing judgment and intuition in addition to providing feedback.

Operational gaming, while just a step away from the traditional model, utilizes role-playing to simulate numerous societal aspects, but could probably be considered a form of simulation.

The general discussion will now concern face-to-face interaction of the type found in most committees and in other decision-making groups. A discussion of the Delphi Technique will then follow.

Chapter 4

FACE-TO-FACE INTERACTION AND DELPHI

In the introductory chapter, it was suggested that there were several well-documented disadvantages inherent in the face-to-face interaction of the usual committee and other group situations. It seems necessary that these be discussed if the advantages of Delphi are to be apparent.

FACE-TO-FACE INTERACTION

"As you know, a committee is something which, when you want a horse, designs a camel" (Dalkey, 14). This may be somewhat on the negative side, but as indicated, there are disadvantages to conducting committees, panels, and other group meetings via face-to-face interaction.

One of the disadvantages is that the result of the discussion is likely to be a compromise between the many divergent views of the members, arrived at, most likely, under the influence of certain psychological factors (Helmer, 30), one of which appears to be the influence of the dominant person. It has been shown that the arrived-at group opinion is likely to be highly influenced, if not totally determined, by the views of that person in the group who does the most talking. Furthermore, it has been found that there is no significant correlation between success in influencing the group and competence in the problem area under discussion (Maier, 39; Kelley, et al., 38; Asch, 2).

Helmer (30) called this factor "specious persuasion", another form of which is evident in the member with the greatest "supposed" authority. That

is, that member to whom the other members of the group ascribe some form of authority, whether or not he actually possesses it. For example, a university committee comprised of a number of department heads and a vice-president could have the opinions of any or all of the department heads disproportionately affected by the stated opinion of the vice-president simply because he was ascribed authority (although here, it is probably actual).

Another disadvantage is the unwillingness on the part of some members to abandon their publicly-expressed opinions, for any number of reasons, all dependent upon the personality of the individual.

Still another, is called the "bandwagon effect" of majority opinion. For example, if it were seen that the majority of the members of a group agreed upon a certain issue, those members holding divergent views could decide to jump on the bandwagon, so to speak, rather than defend their views. This effect is usually noticed in elections, where the candidate whose victory seems to be indicated may, at times, gain support, and at other times, his rival may do so (the "underdog effect") (Rescher, 44).

Another disadvantage is that of the effect of "noise"; that is, redundant or irrelevant material which serves nothing more than to obscure and thus have a biasing effect on that material which is directly relevant to the topic under discussion (Dalkey, 11). Many times, it seems, committees and other groups spend disproportionate amounts of time discussing procedures and group processes as well as other maintenance functions, which may appear to be problem oriented, and very little time remains to thoroughly discuss the issue at hand (Dalkey, 14).

Still another biasing effect is group pressure for conformity. In this case, a timid group member, no matter what level his competence, may have little or no influence on the group's final opinion (Dalkey, 14).

As Dalkey and Helmer (10) explained it:

Direct confrontation...all too often induces the hasty formulation of preconceived notions, an inclination to close one's mind to novel ideas, a tendency to defend a stand once taken, or alternatively and sometimes alternately, a predisposition to be swayed by persuasively stated opinions of others.

In fact, some recent experiments at the RAND Corporation seemed to indicate that, at least when opinions were involved, face-to-face interaction resulted in a group opinion which was less accurate than simply the average of the individual opinions without the discussion (Dalkey, 14).

To conclude on a positive note, though, regarding group interaction, the following statement by Dalkey (14) seems appropriate:

Group interaction can perform a desirable function in the transfer of information, when what is being transferred is knowledge; the problems arise when what is being transferred is opinion.

The discussion will now concern the central issue inferred, and at times mentioned, in the preceding chapters: Delphi.

DELPHI

A series of studies have been conducted, intermittently, at the RAND Corporation since its early days, which were concerned with the problem of the more effective utilization of information derived from group discussions. The early studies were primarily concerned with improvements in the statistical treatment of individual opinions within the context of social and technological events (Kaplan, et al., 36). These indicated that some of the formal properties of individual estimates, such as precision and definiteness, could be used to rate the success of short-term predictions concerning the future. It was also found that background information, as measured by a standard achievement test, had a small, yet significant, influence on the success of predictions.

However, it was found that both of these effects were neutralized by combining estimates into group predictions (Kaplan, et al., 36; Dalkey, 15).

Around 1947, Olaf Helmer and Norman Dalkey of the RAND Corporation in Santa Monica, California, suggested that it might be possible to get a better concept of what the future would be like if one were to ask a panel of experts to give their opinions about it. This idea seemed inherently logical, but attracted little attention at the time (Hayden, 25).

Then, in 1953, Helmer and Dalkey introduced an additional feature into the studies; specifically, iteration with controlled feedback (Dalkey and Helmer, 10; Dalkey, 15). This seems to be the earliest indicated use of the complete technique that was later to become known as Delphi.

The importance of the Delphi Technique, which attempts to make effective use of informed intuitive judgments, is derived from the realization that projections concerning the future are mainly based on individuals' personal expectations, rather than from predictions derived from theory (Helmer, 30).

Delphi attempts to improve the panel or committee approach by subjecting the views of individual experts to each other's criticism in such a way that face-to-face confrontation is avoided and anonymity of opinion provided as well as anonymity of arguments advanced in defense of those opinions (Brown, 4). In essence, Quade (41) called it, "...a 'framework' that replaces a representative model."

The first step in any application of the Delphi Technique is the selection of a group of experts. Decision makers seem to have always depended upon the advice of experts but often this consultation has been haphazard and no attempt has been made to collate differences of opinion among the experts. The selection is still an intricate problem, even with a well-defined category of expertise. A man's expertness might be judged by his status among his peers,

or by years of professional experience, or by his own self-appraisal of relative competence in certain areas, or by the amount of relevant information to which he has access or by some combination of objective indices (Brown, 4).

The discussion of the disadvantages of face-to-face interaction was presented in the first part of the present chapter. Delphi attempts to overcome these disadvantages by utilizing the following three features: 1) anonymity, 2) controlled feedback, and 3) statistical group response. Anonymity is maintained by eliciting separate and private answers to prepared questions, while all other interactions between respondents is only through formal communication channels which are controlled by the experimenters (Dalkey, 11). This serves to reduce the effect of dominant individuals. Controlled feedback, in a Delphi exercise, usually consists of several iterations where the results of the previous iteration are summarized and "fed back" to the respondents. This is a device for reducing the effect of semantic "noise" and other types of discussion not directly problem-oriented. Some form of statistical group response is reported as representative of the group opinion. For those cases where the group task is to estimate a numerical quantity, the median of the individual estimates seems to be the most useful statistical index. No particular attempt is made to arrive at complete unanimity among the respondents, and a spread of opinions on the final round is a typical outcome. This serves to reduce group pressure for conformity and assures that the opinion of every member of the group is represented in the final response ^{there are critics of this} (Dalkey, 11; 15).

A typical Delphi exercise is initiated by a questionnaire requesting estimates of a set of numerical quantities, such as dates for the realization of technological possibilities, or their probabilities of realization by given dates, performance levels, etc. The results of the first round are summarized using the median and inter-quartile range of the responses, and are fed-back

with a request to revise the first estimates where appropriate. On succeeding rounds, those individuals whose answers deviate markedly from the median (that is, lie outside the inter-quartile range) are requested to justify their estimates. These justifications are summarized, fed back, and any counter-arguments are elicited. These, in turn, are fed back and additional revisions elicited. After this final round, a respondent whose answer still remained outside the inter-quartile range was required to state why he was unpersuaded by the opposing arguments. Of course, there can be many variations of this basic pattern (Dalkey, 11; Helmer, 30).

Both the inquiry into and subsequent feedback of the reasons for the opinions that lie outside the inter-quartile range, seems to stimulate the respondents into taking into due account considerations they might, through inadvertence, have neglected, and to give due weight to factors they were initially inclined to dismiss as unimportant (Helmer, 31). Also, forcing the respondents to justify relatively extreme responses appears to cause those without strong convictions to move their estimates closer to the median, while those who felt they had a good argument for a different opinion tended to retain and defend their original estimates (Helmer, 30).

A Delphi exercise has several properties that should be noted. Above all, the procedure is "...a rapid and relatively efficient way to 'cream the tops of the heads' of a group of knowledgeable people" (Dalkey, 15). Generally, it involves much less effort for one to respond to a well-designed questionnaire than, for example, to participate in a conference or write a paper. Properly managed, a Delphi exercise can provide a highly motivating environment for the respondents. The feedback can be novel and interesting if the group of experts involved is mutually self-respecting. From his own experience, Dalkey (15) concluded:

problem then feel no responsibility towards concern

The use of systematic procedures lends an air of objectivity to the outcomes that may or may not be spurious, but which is at least reassuring. ...anonymity and group response allow a sharing of responsibility that is refreshing and that releases from the respondents inhibitions. ...the results...are subject to greater acceptance on the part of the group than are the consensus arrived at by more direct forms of interaction.

One additional feature of recent Delphi procedures should be mentioned. That is, respondents are requested to make some form of self-rating with respect to the questions. Several different kinds of self-ratings have been tried, from ranking the questions in the order of the respondent's judgment as to his competence to answer them, to furnishing an absolute estimate of the respondent's confidence in his answer and estimating a relative self-confidence with respect to some reference group. However, there seems to be no significant correlation between such self-ratings and individual performance in the studies with confirmable estimates. On the other hand, it appears to be possible to use the self-ratings to select a sub-group of relatively more confident individuals whose performance has been slightly, but consistently better than that of the group as a whole (Dalkey, 11). Also, when using the notion of reliability to refer to the range of estimates, the evidence seems to indicate that such systematic processing of expert opinion can produce significant improvements in the estimates, both in accuracy and reliability.

In the majority of cases where the Delphi Technique has been used, a convergence of opinions has occurred. At times, however, no convergence has taken place, but the opinions have been seen to polarize around two distinct values, thus revealing two schools of thought regarding a particular topic. This may indicate that opinions are based on different sets of data or on different interpretations of the same data. If such is the case, then it is conceivable that a continuation of the process through several more rounds might eventually track down and eliminate the basic cause of disagreement, thus lead-

ing to a true consensus. However, even if this does not occur, it should be realized that the Delphi Technique will still serve to crystallize the reasoning process that might lead to one or several positions on an issue, and thus aid in clarification of the topic even in the absence of a group consensus (Helmer, 35).

Whether or not this convergence of opinion favorably compares with that obtained by more traditional modes of consensus formation, is debatable. However, even if the effectiveness of the Delphi Technique in producing consensus is not superior to other methods, it can conceivably offer considerable advantages in cost and reliability; by avoiding the need for assembling the experts in one place, and by not subjecting them to the persuasive nature of oratory or to the "bandwagon effect", but only to the milder form of anonymous social pressure exerted by the feedback of some information on the range of opinions held by the group (Gordon and Helmer, 23).

It may be added here, that when opinions, in actual experiments, changed from one round to the next, the effect was almost entirely explained in terms of the distance of the changed opinions from the group median. The farther away they were, the more likely the opinions would be changed. This is a means of quantifying the effect of group pressure (Dalkey, 14). It was also suspected that the person who resisted such group pressure was more sure of his answer than one who changed. In comparing answers on the first round, it was found that the median answer of those who did not change their opinions (the "holdouts") on the second round, was usually more accurate than that for those who did change (the "swingers"), and actually, more accurate than the median for the entire group. However, when using data from the second round, after changes were made, the holdouts were not as accurate as the entire group, although the swingers were still less accurate. Dalkey (14) described the sit-

uation thus:

The group median was one "center of attraction" for changes, but also the true answer was a weaker center of attraction. Those who changed their mind had a residual amount of information that had not been exploited in the first round, and this improved the overall group answer. In this case, the less knowledgeable members of the group still played a vital role in improving the group response....

In relation to the previous discussion concerning simulation and other techniques, Dalkey (11) stated:

A common reaction is to imagine Delphi as a method of obtaining inputs for some kind of formal estimating structure--e.g., inputs for a simulation model. [however]...for those areas where data is lacking, a formal model is lacking as well. As a matter of fact, the Delphi procedure is one of the most efficient...[known]...for "uncovering" the implicit models that lie behind opinions in the "soft" areas.

SUMMARY

The present chapter has dealt with the problems of face-to-face interaction and the advantages of Delphi in overcoming some of these disadvantages.

It was seen that some of the disadvantages of face-to-face interaction are, 1) influence of dominant individuals, 2) unwillingness to abandon expressed opinions, 3) the "bandwagon effect", 4) effect of semantic "noise", and 5) group pressure for conformity.

It was suggested that Delphi attempts to overcome these problems by the three characteristic features of the process: 1) anonymity of group response, 2) controlled feedback of individual and group opinions, and 3) the utilization of a statistical definition of the group response.

The outline on the following page by Rodgers and Lipsetz ⁴⁵(46) is an excellent representation, round-by-round, of a typical Delphi exercise, and should serve as a useful summary of the process, and a future reference point.

A short discussion then followed, concerning the convergence of opin-

The Delphi Procedure

1. Round I--Round I is an open-ended questionnaire calling for a list of goals, opinions or judgments based on one's experience in the organization or field being studied.

The open-ended question should be comprehensive enough to solicit a very wide range of comments dealing with the task for which you are gathering data.

2. The comments from the Round I question must be categorized and sub-categorized until items can be written that are truly representative of the data received. While there is no standard way of building these items it is best that more than one person be involved and that the initial categorizing be done separately. It is important that no items be written that are not represented by the initial data received.
3. Round II--Round II is sent to each participant in the Delphi. They are asked to rate or evaluate each of the items by some criteria such as rank or importance, etc.
4. Round III--Round III repeats the same items as in Round II. The participant is given his response to each item from Round II as well as information as to how the rest of the group responded to each item. An example would be the modal response of the group for each item. The participant is asked to respond to each question once again. Instructions tell the respondent to answer with the group response for that item unless he has a specific reason to do otherwise. If he answers outside the group's response for that item he is to write his reason in the margin opposite the item.
5. Round IV--Round IV repeats the Delphi items for a third time. This time the participant is given his response for each item for Round III as well as the consensus of the group for each item from Round III.

Also, the minority comments from Round III are provided. Again he is asked to answer within the group consensus unless he has a specific reason to do otherwise. If he does have a minority opinion he is asked to list it in the margin provided.

ion, and the effects of group pressure on accuracy of answers where results could be confirmed. The discussion will now turn to a review of Delphi in certain applications and in actual experiments, with suggested refinements.

Chapter 5

APPLICATIONS AND REFINEMENTS OF DELPHI

APPLICATIONS

As suggested in the introductory chapter, the first indication of the actual development and use of the Delphi Technique seems to have occurred around 1953. However, in 1950, a report was published which purported to study the "prediction of social and technological events" (Kaplan, et al., 36). It was a pilot study, primarily concerned with the following three problems with respect to prediction: 1) evaluation, 2) improvement and 3) appraisal.

The questionnaires used were evenly divided in subject matter between social science and natural science events, and the predictions concerning the events were of the form "the likelihood that E will occur is p" (categorical) rather than "if the conditions C are satisfied the likelihood of E is p" (conditional), and concerned future events on which the predictors could have little or no effect.

Four exhaustive and exclusive alternative outcomes were offered for each question, with a time limit for occurrence of the event set at 20 weeks (or less) from the date of the questionnaire. This study, therefore, can be classified as one involving short-range predictions.

With respect to the problems investigated, the experimenter arrived at the following conclusions: Regarding evaluation, it was found that individual predictors could be evaluated in terms of success, definiteness, and precision; that is, frequency of correct responses, how definitely they were

committed to a prediction, and the accuracy of their estimates of the probabilities of the given alternatives. But while these aspects seem to be associated to some degree, the results did not appear to indicate that they could all be regarded as resulting from some single skill in prediction.

It was also concluded that the relatively near future seems to be more predictable than the relatively distant future. In addition, as to the question that some types of events may be easier to predict than others, it was found that the data suggested a somewhat greater predictability of the social science, as against natural science, subject matter.

With respect to the problem of improvement, it was found that the success of the best informed predictors was not that much greater than the success of the worst informed, although it was in the positive direction. Also, there were no significant differences in the relative frequency of success with questions of varying degrees of difficulty.

A special phase of the study was directed at studying the effect of having the prediction made by several predictors together, as a suggested method for improving the success of predictions. It was found that this group effort was significantly better than that of the same individuals working independently. However, it was questioned whether this result was due to collective effort on the part of the group, or whether it was due to a certain averaging that could equally be obtained by a statistical combination of the individual results. It was found that this was indeed the case, in that the result of the group effort was statistically duplicated.

With respect to appraisal, the data showed that knowledge of the rank order of likelihood assigned to the predictive alternatives would verify the first choice among alternatives more often than it would the second choice, and the second choice, in turn, more often than the third. In addition, the level

of likelihood was studied, and in regard to both aspects, it was found that both the rank and the level of a prediction provide a useful basis for the appraisal of the chances for the verification of the prediction.

The study also examined another way of appraising a prediction, that is, the grounds on which the predictor says his prediction is based. It was concluded that this factor can discriminate sub-populations of predictions which differ from one another in success, thus improving the estimate of reliability of prediction, by content analysis of the basis statements made by non-experts (Kaplan, et al., 36).

As can be seen, the above study does not represent a complete Delphi, the only things it shares are the use of questionnaires and some experts, and some indication that groups can improve the estimates obtained. The earliest indication of a study that seems to approximate the many facets of Delphi, and indeed, used a process that was, at the time, called Delphi, was reported in 1963, in an article by the inventors of Delphi, Norman Dalkey and Olaf Helmer (10). The study was actually conducted around 1953, the content of which, for security reasons, was not released until 1962. This seems to be the same study in which Dalkey and Helmer first introduced their notion of iteration with controlled feedback as alluded in the preceding chapters.

Dalkey and Helmer provide a good brief description of the intent and method of the study in the following statement:

In this application, expert opinion was applied to the dual problem of the selection of an optimal industrial target system and the estimation of the number of A-bombs required to reduce the munitions output by a prescribed amount. Seven experts participated, responding to five questionnaires submitted at approximately weekly intervals. The first questionnaire was followed by an interview in which each respondent was asked to reproduce the reasoning by which he arrived at an estimate of the number of bombs and to show the component breakdown by industries. The third also was followed by an interview for the clarification of ambiguities....

The information fed back to the experts between rounds was generally of two kinds: 1) available data previously requested by one of the experts, or 2) factors and considerations suggested as potentially relevant by one or another respondent. Although the experimenters did not expect to achieve absolute consensus, they did expect the opinions to converge, due to "the progressively more penetrating analysis of the problem", by the nature of the process itself. The numerical quantity being estimated did show considerable convergence, with the ratio between the largest and smallest response dropping from 100 - 1 on the first round, to 3 - 1 on the final round (Brown, 4). In total, five questionnaires were used, thus this is a 5-round Delphi exercise, and as will be noted later, more recent ones do not go this far, due to the evidence suggesting that convergence is minimal and little is gained beyond the third round.

The experimenters seemed conscious of some of the limitations of their process, and indicated so. First, the responses of the experts were not considered strictly independent, as some working assignments required contact between several respondents. Also, the experimenters used at least one respondent as a consultant on one aspect of the subject matter. Some "leading" by the experimenters also resulted from the selection of the information supplied by the experts (Dalkey and Helmer, 10).

The first application of the Delphi Technique to long-range forecasting was made in 1964 (Gordon and Helmer, 23) in an investigation in which several panels of experts were asked to make contingency forecasts of the state of the world twenty-five to fifty years in the future. Six groups of experts were selected, one group for each of the following six areas of inquiry: Scientific breakthroughs, population growth, automation, space progress, war prevention and future weapon systems. Each panel answered four sequential ques-

tionnaires spaced approximately two months apart and made judgments on more than two hundred predictive items during the course of the experiment (Brown, 4).

As Dalkey (15) stated, in reference to long-range forecasting:

In the area of long-range forecasting, it is difficult to dodge the fact that a large part of the activity is at least within the area of opinion, and possibly worse. That particular study [Gordon and Helmer, 23] happened to coincide with a surge of interest in long-range forecasting itself, with an attendant interest in the systematic use of expert opinion.

Each group in the study was asked questions specifically oriented to their particular area. In the first questionnaire, they were asked to list major events that might occur in the next fifty years. These items were then compiled and returned to the panel via the second questionnaire, which was a request to estimate the time of occurrence of each event. The items from this questionnaire which were observed to possess a "reasonable concurrence of opinions" were so listed on the third questionnaire. Those items that remained were repeated for further consideration, along with the range of opinions received and a request for those maintaining an extreme position compared to the majority to state their reasons for such a position. In the fourth questionnaire, the opposing statements were presented for consideration, and a request was made to arrive at a final, perhaps revised estimate of the occurrence of each event. The result was then expressed in terms of the median of these final responses together with their quartiles (which indicated the spread of responses and thus, the degree of consensus achieved) (Helmer, 27).

Specific results from this study, in the area of methodology, were lacking, due to the absence of full experimental control; and due to the nature of the questions, there could be no verification of the validity and reliability of the forecasts. However, it was concluded that some objective

means of discriminating among respondents concerning their competence seemed desirable. The following experiment was conducted to investigate this (Brown and Helmer, 3).

The questions submitted were of the "almanac" type for which answers were available, thus the validity and reliability of the estimates could be checked. Using a four-round Delphi process, an extra item was included: the respondents were asked to indicate a self-rating on a scale from 1 - 4, which would be an evaluation of their own degree of expertise on each question. This was done on the opening round and on the final round. Upon evaluating these self-ratings, a sub-group was selected comprised of those individuals who gave themselves the highest ratings, about one-third of the total group, for each question.

The median for this group on the fourth round was compared to that of the entire group, and a considerable improvement in the estimates was noted. Thus, it was concluded that this indication of competence was a powerful tool for increasing the reliability of the group estimates (Helmer, 27). Also, in checking the median answers obtained in the study, with the actual answers, a convergence of the medians toward the true values occurred in most of the cases. Helmer (27) commented on the use of the median to define consensus:

If a single number is to be used to be representative of the group's collective opinion, the median has the evident advantage over...the mean of being independent of the scale. Moreover, it has the obvious and intuitively appealing quality that it is that value for which half the group thinks that the true answer is less than or equal to it and the other half that it is greater than or equal to it....

Brown (4) summarized a report which used the Delphi Technique to study forecasts of business and economic indices. Four groups were asked to make forecasts of sixteen economic series for the first quarter of 1966. Two groups used the traditional methods of making business forecasts and the other

two used the Delphi process. The traditional method allowed free interaction among participants in the group for the purpose of obtaining information relative to the forecasts. The Delphi group responded to a series of four questionnaires over a period of six weeks. The group that used the Delphi process was found to have made more accurate forecasts than the group using the traditional business forecasting technique. However, the meaning of "more accurate" was not clarified.

This has been a review of some of the very early experiments utilizing actual Delphi procedures and parts thereof. The following discussion will concern more recent applications of the technique in reference to several areas. Specifically, these areas are: 1) values and evaluation, 2) quality of life, and 3) the future. In addition, a discussion of some suggested refinements of the process will follow.

Values and Evaluation

In formulating policies and in making decisions, two different kinds of inputs seem to be involved. One is factual judgment, and the other is value judgment. The experimental work on Delphi procedures has usually dealt exclusively with factual judgments. However, some applications of the procedures have attempted to elicit and process value judgments, a fairly popular form of which is the formulation of the major objectives of an organization and the weighing of these objectives on some scale (Dalkey, 15).

Rescher (42) agreed with Dalkey in considering Delphi for inquiring into the area of values, and his report was aimed at outlining some theoretical background considerations for applications of Delphi procedures in the value area. Rescher (42) noted that,

When one steps outside the traditional area of application of Delphi to factual issues, one perhaps arrives first of all at

the possibility of its use as an instrument for decision-making. Just as we can use Delphi in the traditional way to explore the prospects of a group consensus regarding "what the facts are (or--in predictive applications--will be)" so we can deploy it on the issue of "what to do". This step does not yet carry us into the value domain--although it takes us to its threshold. However, we do not cross that threshold until we inject an overtly evaluative element of "what ought to be done" with its reference to the positive and negative, the favored and disfavored aspects of the case.

A report by Helmer (28) concerned the design of a five-hour simulation workshop on the subject of the effect of technological change on American values in which two operations research techniques were to be employed: simulation and the systematic use of expert judgment.

The subjects were to be involved in a simulated planning process in which they would be asked, 1) to make decisions which would affect the character of the environment, 2) to estimate the societal consequences of these decisions, and 3) to evaluate the desirability of the consequences. They would not be asked to abstractly judge values, but to state their preferences when presented with a simulated decision-making situation. "Thus we shall be dealing not with future values directly but with potential future environments against which a future values structure is to be judged" (Helmer, 28). In actuality, the procedures were considered to simulate a simulation of a social decision-making process in that the subjects were to be asked to take a critical look at the method of simulating the process of social planning as a means of arriving at a moral evaluation of alternative futures. There are no results to be reported concerning this workshop, as this was a report of the design of the workshop, rather than of its actual running.

Dimon's original Rescher (43) felt that most of man's actions reflect his predisposition to make choices. These choices manifest his preferences which, in turn, mirror his values. The technological and social environment in the future will, therefore, be the reflection of his future values. The scientific study

of values is recent and still very much underdeveloped, and predictive instrumentalities for such study are presently lacking.

One obviously appropriate application of the Delphi method is for discovering the values of a group. Delphi is a tool for exploring group opinion, but there seems to be no reason why it should not be directed at group opinion about values as well as facts. At this level, Delphi could provide a means of explaining the structure of group values by focusing not only on choices, but upon the reasons for making choices one way rather than another, and upon the weights of these reasons, thus revealing which values are regarded as relevant to the discussion (Rescher, 42).

One experiment was reported that attempted to study the use of Delphi in the area of evaluation, and it will now be briefly reviewed.

The report by Dalkey and Rourke (18) described the results of an experiment assessing the appropriateness of Delphi procedures for formulating group value judgments. Three factors were investigated: 1) the quality of the distributions of the responses, 2) the correlation between ratings by different rating techniques, and 3) the amount of change and degree of convergence upon iteration with feedback.

The results of the study showed that the distributions, as expected, were "normalized". The correlations between different groups and different rating methods were high (.90 or higher), and the number of changes and degree of convergence (reduction in standard deviation) were both comparable to similar indices obtained from experiments using factual judgments. It was concluded that Delphi procedures were appropriate for generating and assessing value, as well as factual material (Dalkey and Rourke, 18). Further, the authors continued, from the standpoint of the decisionmaker, it seems that opinions about values and objectives are just as relevant to decisions as fac-

tual opinions about consequences. However, there seems to be no way to measure the correctness of value judgments; there is no generally agreed-upon set of "facts" against which they can be compared. Furthermore, value judgments seem to be "emotionally loaded", in that the expression of value judgments appears to be more directly tied to emotions than are factual statements, and commitment to those judgments may be more central to the personality of the individual, so that the interaction of value judgments with other cognitive material (such as facts) is impeded (Dalkey and Rourke, 18).

Rescher (42) agreed:

...in one basic respect there does seem to be a...limitation in the use of Delphi in the value area...there is no generally practicable technique for checking the value judgments of the group against "the actual facts".

It appears to be the case, then, in disagreements about values, that most people would state that one side can be more correct than the other without being able to specify how the value judgments can be validated. "As it turns out, it is not necessary to be able to specify what correctness or incorrectness means in order to say a great deal about better and worse judgments" (Dalkey and Rourke, 18).

Quality of Life

The phrase "quality of life", (hereby referred to as "QOL") seems to have supplanted the older words "happiness" and "welfare" in contemporary discussions of urban and domestic policy (Dalkey, 12). However, few users of the phrase bother to define it. Still, some have attempted to give content to the notion and these attempts seem to have taken two forms, 1) armchair "analyses", and 2) public surveys.

The armchair approach usually consists of devising a list of general factors which are presumed to be significant in determining the well-being of

people. The lists may consist of clinical lore, sociological concepts or results from psychological and social-psychological experimentation, and there appears to be a great deal of overlap in that the shorter lists obtained from this approach are usually contained, bodily, in the longer ones (Dalkey, 12).

A preliminary Delphi exercise was conducted by Dalkey (11) using twelve RAND staff as a panel. The panel was asked to consider the following nine factors: health, activity, freedom, security, novelty, status, sociality, affluence, and aggression. They were asked to determine the relative weight of the factors for the quality of life of the average American, as well as to determine whether they were meaningful and/or measurable. The panel was also asked to add any new factors which they thought were significant.

The results of the exercise showed that almost everyone agreed that the items were meaningful, and that all were measurable except for freedom, novelty, and aggression. Considerable diversity on the values of the relative weights was encountered, but reasonable agreement on their ranking was achieved (Dalkey, 12). It was concluded that QOL seems to be determined mainly by some very general features of the individual and his environment, rather than by specific factors. Largely, the issue is whether what is being sought is a "single thing" that may be called QOL, or whether it is presumed to consist of a number of incomparable elements. In addition, there seem to be two other propositions for the determination of QOL, for which there is a fair amount of evidence. First, it appears to be the case that the influence of factors on QOL are a rapidly decreasing function of distance away from the individual, either in space or time. Second, is that human beings seem to live much more "in the future" than lower animals, thus such things as hope, anticipation, ambition, aspiration level, and anxiety seem to be important elements of QOL. Indeed, it does appear reasonable to assume that events of the

distant future are much less influential than near events.

The second of the means for attempting to give form to the notion of QOL is via public surveys. These are usually based on lengthy interviews involving over two hours and over one-hundred items. Questions may range from the subjective and global to the objective and specific, however, there are many obvious problems with this approach, only one of which is the reliance on questionable verbal reports (Dalkey, 12).

Dalkey (12) presented a research proposal which would attempt to be more systematic in studying the quality of life than the other two techniques described above. In summary, the basic idea was to "...prepare a comprehensive set of scales relevant to the quality of life; let a large, representative sample of Americans rate themselves...; and employ factor analysis to summarize the interrelations between the ratings". There would be three sorts of scales: 1) objective measures (such as income, age, etc.), 2) subjective ratings (such as job satisfaction, perceived social status, etc.), and 3) global subjective items (such as happiness, optimism about the future, etc.).

The author saw one of the difficulties as the inclusion of items on such things as aggression, anti-social behavior, bigotry, etc., and stated: "The presumption that the quality of life is determined solely by 'acceptable' items is, of course, false", and felt that these difficult areas would probably have to be underemphasized on the first sequence of questions (Dalkey, 12).

Furthermore, the most critical part of such a study, and the one that could probably consume a great deal of the elapsed time, is the construction and selection of the set of primary scales, since their potential is almost limitless. One of the problems of such a study, as with public surveys, is the reliance on verbal reports, which have, as indicated, questionable aspects. In addition, a few other internal problems would have to be considered before

the study could be implemented.

Dalkey (12) also saw several reasons for such a study, but singled out the use of the factor analytic approach, which would provide a systematic framework for tying together a vast amount of information about the perceived well-being of Americans.

Another report (Dalkey, et al., 17) concerned a preliminary model for the analysis of the quality of life, much along the same lines as the study suggested by Dalkey (12), and provided some illustrations of its applications to choices in career and transportation. The model was considered useful in suggesting lines of investigation and in interpreting data. It included a set of general characteristics or qualities of the stream of events occurring to an individual that largely determine his sense of well-being. It was mentioned that several Delphi studies which were guided by the model produced similar lists of qualities in which the ratings of importance were comparable. However, the authors considered the Delphi Technique as probably not too well suited to the establishment of ultimate results, but saw its usefulness in exploratory research and in uncovering hypotheses that can be confirmed by other methods of investigation such as cross-sectional surveys or behavioral experiments. This report was so extensive and detailed in its discussion of theoretical and philosophical issues involved in the study of the quality of life, that an attempt at further synopsis would carry the discussion beyond the range of the present paper. However, the report is considered an excellent point of departure for those interested in studying the quality of life.

The Future

As mentioned earlier, Delphi was initially designed as a device for polling experts about their predictions--especially about the future--as im-

plied by the word "predict". It seems appropriate, at this time, to provide a review of the concept of the future as encountered in the literature on Delphi and comments regarding the utilization of Delphi and Delphi-like procedures in the study of the future. A statement by Hayden (25) will serve as an introduction to this discussion.

Today you can open the newspaper almost any day and read some startling prediction about the future. The future has become so much a part of our present that it has become respectable to speculate on almost any trend or tendency. Scholars of great eminence, supported both by the foundations and by the government, are seriously engaged in attempts to foresee the future.

good keep! There seem to be two good reasons why it is important to look into the future: 1) to identify things that are presently being done that must be stopped if man is to survive, and 2) to determine the consequences of some of the things that may be under consideration or that may occur accidentally, by projecting as far into the future as possible.

One can make wild conjectures by drawing most trend lines far enough into the future, whether they are crime statistics or divorce rates. However, circumstances change and judgment must be used in projecting trends. Furthermore, interactions between competing futures must be sought and decisions made with respect to the actions that are taken to affect the outcomes. "Decisions being made today directly affect the future. Some of today's decisions have a statistical probability of affecting someone we have never seen in distant lands, and may never see--in future generations" (Hayden, 25).

Helmer (30) felt that, beginning with the 1960's, there has been an important change in the intellectual climate throughout many parts of the world. This change was evidenced by a new attitude toward the future that became apparent in planning agencies as well as in the research community. The effect of this change has been to extend customary planning horizons into

a more distant future and "to replace haphazard intuitive gambles, as a basis for planning, by sober and craftsmanlike analysis of the opportunities the future has to offer." Helmer continued:

The change in attitude toward the future is manifesting itself in several ways; philosophically, in that there is a new understanding of what it means to talk about the future; pragmatically, in that there is a growing recognition that it is important to do something about the future; and methodologically, in that there are new and more effective ways of in fact doing something about the future.

Helmer elaborated on each of these three aspects. Philosophically speaking, the exploration of the future is no longer equated with fortune-telling or crystal-ball gazing. Instead, there is a growing awareness that much can be said about future trends in terms of probability, and that through proper planning, these probabilities can be considerably influenced. "The future is no longer viewed as unique, unforeseeable, and inevitable; there are, instead, a multitude of possible futures, with associated probabilities that can be estimated and, to some extent, manipulated."

Pragmatically speaking, not only are technology and the environment undergoing change, but the pace of change is constantly accelerating. It is becoming mandatory to strive to anticipate changes in the environment rather than attempting to deal with them belatedly and inadequately after they have become obvious.

Methodologically speaking, the use of operations research techniques, in the form of mathematical models, simulation procedures, and the systematic use of experts, to study the future is becoming increasingly widespread. Plus, new uses of computers, with automated access to central data banks will provide massive data-processing capabilities.

Rescher (43) discussed the forecasting of future changes in human affairs in the scientific and technological areas and human and social environ-

ments. He made a statement that seemed intuitively logical, "...it is only natural that one should be interested in the future, since we're all going to spend the rest of our lives there."

Rescher further contended that the key to the new attitude toward the future lies in the concept of planning. Planning is a concept, the importance of which, is recognized in departments of government, education, research, industry, and labor, to mention just a few. The value of planning has been evidenced by the problems posed by the strains on society of the phenomena of economic fluctuations, automation, educational inequities, urban congestion, pollution of air and water, all of which could be removed or relieved by the use of foresight in the planning for preventive measures. "The premium put on planning by the increasingly high price that must--in the current context--be paid for by the traditional policy of meeting difficulties as they arise and 'muddling through' has given a new-found respectability to future-oriented studies."

Rescher (43) also suggested that another important task for the development of an adequate predictive instrument for studies oriented to the future would be the systematic study of the formation of intellectual and social fashions. Neither the desirability nor the actuality of extensive researches into the future seems questionable at this time. However, the scientific standard of such investigations represents a still unsettled issue.

Dalkey (14) stated it another way:

Science, technology, and advanced administrative methods are rapidly extending the scope of public and private control over economic and social developments. To use these new powers wisely, it is necessary to have a clearer picture of the future--or more exactly of the possible futures.

The improved understanding between experts and the public is the goal of a new, computer-based "game" developed at the University of Illinois. This

game, called Delphi, involves people as "explorers" of possible future developments. During the exploration they are presented with information about possible events and then are asked to indicate how they would like to change the probability of the occurrence of each of the events. However, this Delphi Exploration described here should not be confused with the Delphi Technique. Both make use of lists of developments or events, and both may employ computers. But the Delphi Technique consults experts for estimates of probability, while the Delphi Exploration consults members of the public for estimates of desirability. The two procedures are considered complementary, rather than competitive (Umpleby and Briggs, 49).

Helmer (34) also expressed his thoughts on another aspect of the study of the future:

The survey of possible futures, with which any analysis of the future must begin, will continue to have to rely primarily on the intuitive judgment of experts. The process of obtaining a consensus among specialists will be enormously improved.... Not only will the expected gigantic increase in scientific knowledge raise the quality of available expertise..., but the day is not too far off when we can establish a world-wide network of specialists, each equipped with a console tied to one central computer and to electronic data banks, who will be able to interact with one another via the computer network and thus obtain a consensus among themselves through a process of which the present-day Delphi technique is a primitive precursor.

Goldsen (22) considered the future, like the past and the present, as essentially man-made. That is, what man does, what he dreams, what he thinks, and what he consciously and unconsciously remembers has constituted the past, accounts for the present and will also shape the future. There are connections between the present but changing perceptions of the past, which are known or knowable, and these reach out with wide continuities into the future. Because of these, it should be easier to describe the future in an "aggregate" sense than it would be to predict discrete technological inventions or to predict

sudden discontinuities, abrupt accelerations or declines in rates of change in the social world, especially if one wants to specify dates, probabilities or magnitudes.

Goldsen (22) also stated:

One of the most difficult tasks in planning for future development is not only to establish the present preferences, value systems and objectives of differential groups but to understand how these may change and to allow for the likelihood that what we think is wanted or desired today may not in fact be so preferred in the future.

*what we
are attempting*

Helmer (30) in speaking on the Delphi Technique for the use of studying the future, explained that numerous further experimentation is needed to test the extent of its validity and "to refine it to the point where it may be fully accepted as one of the standard tools for the analysis of the future and ...for policy applications in the general area of social technology."

An experiment, conducted to investigate long-range forecasting, was described earlier (Chapter 5, p. 31) and will not be reviewed again, even though its implications for the study of the future seem obvious. However, the results of the experiment did illuminate a number of points: the content of the predictions, the respondents' bases for their predictions, the spread of the views of the experts, the convergence of those views following data feedback, the experts' critiques of each other's views, as well as the weaknesses of the method and the possible means for improving it. Unfortunately, the experiment was so detailed with respect to these points, that an adequate summary would prove not only prohibitively long, but would perhaps remove the present discussion from its desired purpose.

However, this is not to say that Delphi does not have its uses in the study of the future. It seems to have had an impact on the study of the future in the educational setting, and this will be discussed in the next chap-

ter, in the section entitled FUTURISTICS.

For the moment, the discussion will turn to some of the suggested refinements for the Delphi Technique encountered in the literature.

REFINEMENTS

A number of variations of the Delphi process seem worthy of exploration. Among them, two refinements have been experimented with and appear to show a good deal of promise.

The first variation introduces the concept of weighted opinions. If it were easy to objectively measure the relative trustworthiness of various experts, the greatest, if not exclusive, weight would have to be given to the opinions of those considered the most trustworthy. In view of the absence of such measures, the idea of relying to some extent on the expert's self-appraisal of their relative competence has been investigated and the results found quite promising. One study utilized this device with twenty members of the UCLA Business Administration faculty who made forecasts of ten economic and business indices. The procedure had the respondents participating in four rounds of Delphi, and also ranking their relative competence regarding each of the ten indices. Then, instead of using the median of the final responses to form the group consensus, only the responses of the individuals who ranked themselves relatively highest for each particular index were used, and the median of these responses was employed to arrive at a group consensus (Helmer, 29).

The statistical evidence for the superiority of this selective use of self-appraised responses seems insufficient; however, in the few cases where it has been tried and subsequently checked against the facts, it has led to a more reliable consensus. In fact, this select median, when compared to the

median of all responses, was closer to the true value in two-thirds of the cases in which it was used (Helmer, 35).

A second refinement of the basic Delphi process is in its application within a simulated decision-making exercise. A typical situation to which this use of expertise is applicable is in making recommendations for budgetary decisions on the basis of cost-benefit estimates. "Delphi procedures, built into the simulated decision-making process, can be used as a systematic, rational approach to what otherwise would be a haphazard compromise effort" (Helmer, 29).

Helmer (35) suggested that further refinements can include the introduction of subsidiary questions and the removal of systematic bias. By subsidiary questions, Helmer referred to the possibility of inviting the respondents, as part of the first-round questionnaire, to suggest subsidiary questions, the answers to which could help them arrive at a more reliable answer to the central question. If these questions were to refer to existing statistical data, for example, they could then be supplied along with the second questionnaire. Otherwise, they could give rise to a subsidiary Delphi inquiry among the experts.

In reference to the removal of systematic bias, Helmer stated:

...it has recently been suggested that there may be a negative correlation between a forecaster's age and the date of occurrence of an event predicted by him. This, and other similar correlations, ought to be looked at, and if found true we would be in a position to correct for such systematic bias.

The RAND Corporation, in recent years, has begun to experiment with the possibility of combining the Delphi process with computers. Specifically, the experiments have been conducted utilizing what is known as JOSS on-line computer consoles which replace the ordinary paper and pencil method of replying to a Delphi questionnaire. By having each of the experts enter his

responses on the console, the group's responses can be automatically processed and information immediately fed back with the instructions that make up the next questionnaire. Thus, rather than taking weeks or months, the entire process can be carried through in an hour or less. Moreover,

...once the interaction among the respondents is via machine, it would be relatively easy to enrich the process by providing on demand automated access to existing data banks and eventually even to banks of mathematical models that might aid the expert in the analysis of the situation under consideration. Once this process has been perfected, it is easy to imagine that for important decisions simultaneous consultation with experts in geographically distant locations via a nation-wide or even a world-wide computer network may become a matter of routine.

A few words might also be said about some modifications of the Delphi Technique that would make it of use in situations where face-to-face interaction is unavoidable or even expedient. In the case of a small committee, for example, some of the inadequacies of the face-to-face encounter (mentioned in Chapter 4) can be removed without going all the way to the full anonymity guaranteed by a complete Delphi procedure. Helmer (35) provided an illustration:

A simplified version of Delphi, which has been found useful...goes something like this: Have each member of the panel independently write down his own estimate, reveal the set of estimates but without identifying which was made by whom, and debate openly the pros and cons of various estimates; then have each person once more independently write down his own (possibly revised) estimate, and accept the median of these as the group's decision.

From this simple case, where the outcome is the determination of only a single number, one can proceed to more sophisticated uses, such as simulated planning exercises or even actual planning. One example, that is simple in concept but requiring a more elaborate procedure than that involved in the above illustration, is that of allocating a given budget among a number of projects. The difficulty, of course, is that an allocation to a particular project cannot be handled in isolation, but that an integrated view of the en-

tire series of projects and their interrelations must be considered (Helmer, 35). (See the discussion of Cross-Impact Matrix forecasting later in this section for a similar rationale).

Another refinement might be the use of a "hierarchical panel structure". This would be a Delphi procedure where the responses are collected, not from individuals, but from sub-groups of experts, and where each sub-group, in turn, arrives at its group response by a Delphi inquiry among its members. However, in (Dalkey, 13), no criteria was found which enabled the selection of a "superior" sub-group which would out-perform the entire group. Attempting to use a self-rating scale to denote confidence or relative performance does not seem to offer a reliable way of singling out such superior sub-groups. The reader may see a contradiction between the above statement and that from Helmer (27) as reported earlier in the present chapter (p. 33). The two differences, however, seem to lie in that here, the selection of the sub-group is made prior to the use of the procedure, whereas it was made after the procedure in the earlier study. In addition, the other sub-groups were determined statistically, and did not exist in actuality, whereas these did actually exist. A direct comparison of the two means of forming sub-groups can not, therefore, be made.

In an attempt at a clearer understanding of the Delphi process, and to lay a more firm foundation for improvements, another experiment was conducted with what was termed "closed information". That is, during the exercise, no new information concerning the subject matter was introduced into the group. But even in this case, the accuracy of the group response increased with iteration--"rather like lifting itself by its logical bootstraps on the part of the group" (Dalkey, 13).

The reader will recall that Helmer (35) mentioned the use of JOSS con-

soles for use with Delphi and that RAND had, at the time, been experimenting with the procedure. This has now become a reality as reported by Turoff (48).

Automation of the Delphi technique by substituting computers for the usual paper, pencils, and mailing of questionnaires results in something that is no longer a Delphi exercise, but a new system with significantly different properties. It is better referred to as a Delphi Conference.

The computer, in a Delphi Conference, stores discussion items entered by group members and accumulates votes on these items. Then, since this is an up-to-the-minute record of the status of the conference, maintained by the computer, an individual at a terminal can have a display summarizing all that has occurred up to that point, and can then enter his views to become part of the record. With this arrangement, a member can contribute to the conference any time of the day or night, and have it recorded, even though he is thousands of miles from another participant; the commands can be transmitted to the computer via ordinary telephone lines.

In fact, the U.S. Office of Emergency Preparation in Washington has utilized this method to explore its own potentialities. In an initial experiment, utilizing twenty respondents located across the country, any one of them could interact with the computer in any of the following ways: 1) he could have the computer display summaries of the conference proceedings, 2) he could have it display a text of a particular discussion item, 3) he could express a judgment or vote on any discussion items not previously judged, or change his earlier votes and judgments, 4) he could add one or more discussion items for group consideration, 5) he could add a commentary message not for evaluation by the group, which could have only a limited life-span on the computer, or 6) he could obtain a summary of his vote for all the items. The discussion items were in the form of comments, proposals and estimates. Comments were evaluated on a primary scale of importance, proposals on a scale of de-

sirability, and estimates in the form of numerical data and a vote on the primary scale of importance (Turoff, 48).

The Delphi Conference makes efficient use of the time of the respondents, and some other uses of the conference include: 1) making sure all issues are on the table and everyone has had an opportunity to consider the issues before a committee actually convenes, 2) maintaining a continuous dialogue over weeks, months, or even years between a group of persons, and 3) to improve lateral communications in an organization. However, this latter use may receive little attention in those organizations still using heirarchical communication structures (Turoff, 48).

Forecasting techniques such as Delphi, based on the collation of expert judgment, suffer from the possibility that interactions between forecasted items may not be fully considered. A new forecasting technique seeks to find the conditional probability of an event given the occurrence or non-occurrence of various other events. That is, potential relationships between the forecasted events may be ignored and the forecasts could then contain either mutually reinforcing or mutually exclusive items. As Gordon and Hayward (24) indicated: "It is hard to imagine an event without a predecessor which made it more or less likely or influenced its form--or one which, after occurring, left no mark." This interrelationship between events and developments is called "Cross-Impact", a term suggested by Olaf Helmer.

The Cross-Impact Matrix method of forecasting is vastly complex, and a systematic description of all potential modes of interaction would be very difficult to summarize. However, two experiments were conducted utilizing the method, and these will be briefly discussed.

The first experiment was termed an "historical test" and utilized a computer. The results were considered consistent with the expectations of the

experimenters. True, this is a very general statement, but when the probability estimates obtained in the study were treated with cross-impact, the events which emerged as being the most probable were "a logically consistent set (though not necessarily more logical than other sets)" (Gordon and Hayward, 24).

The second experiment was a "methodological test", where a cross-impact matrix was used with 71 developments and events relevant to transportation within the next twenty years, forecasted with varying levels of probability. A matrix was constructed after the dates of occurrence and initial probabilities for each item were estimated; it showed modes and strength of linkage, and predecessor-successor relationships. The results of the matrix were assessed in terms of the amount of shift in probability and final probability. As Gordon and Hayward (24) concluded:

It appears that modes of linkage between events and developments can be grouped into several categories, each with its own properties. This generalization permits at least a primitive analysis of the potential interactions between the items. In the two cases examined, this analysis has led to some insight about the future which was not available by inspection of the items alone; it came when the interactions between the items were explored.

One final note is worth mentioning, and that is, that it seems that the orderliness of the matrix tends to force an investigator to be very explicit about the relationships he believes to be functioning in his field of study.

On the following page the reader will find a sample illustration of a cross-impact matrix arrangement for a small number of developments. It is the same illustration found on page 102 of the Gordon and Hayward (24) report.

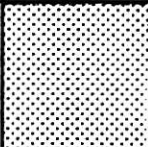
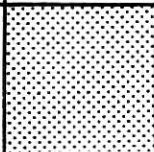
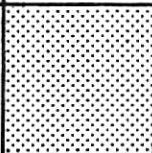
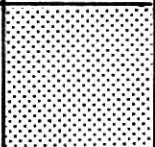
SUMMARY

The present chapter has attempted to review some of the applications and refinements of the Delphi Technique. The technique was noted to have been

By way of illustration, if the following developments and probabilities were forecast for a given year:

Development D	Probability p
D_1 One month reliable weather forecasts	0.4
D_2 Feasibility of limited weather control	0.2
D_3 General biochemical immunization	0.5
D_4 Crop damage from adverse weather eliminated	0.5

then these might be arranged in matrix form as follows:

If this development were to occur:	then probability of			
	D_1	D_2	D_3	D_4
D_1		—	—	↑
D_2	↑		—	↑
D_3	—	—		—
D_4	—	—	—	

where the "up" arrows indicate positive cross impact.

Thus if D_2 ...were to occur, D_1 ...and D_4 ...would become more probable as noted by the upward arrows.

developed around 1953, and has since been utilized in a number of various ways.

The uses or implications of Delphi in the study of Values and the process of evaluation was reviewed, and a limitation to its use in this area was noted--there are no facts against which to check the value judgments obtained by a Delphi procedure.

The study of the quality of life was reviewed and it was seen that there were two approaches, the armchair approach of devising a list, and public surveys involving long hours and hundreds of questions. A preliminary model for such a study was briefly reviewed, and it was seen that several Delphi studies which were guided by the model produced similar lists of qualities. Delphi's use here seemed limited to exploratory research and in uncovering confirmable hypotheses.

The study of the future was then reviewed, and it was mentioned that during the 1960's, the future began to be considered as an object of research, capable of manipulation, rather than an unknowable quantity. Delphi was considered a primitive form of what could someday be a world-wide computer network providing valuable information to specialists. It was noted that Delphi has already been combined with computers, and that a derivation of the concept, using members of the public rather than experts for the study of the future, called the Delphi Exploration, has been developed.

Some refinements of the technique were noted, and that such things as weighted opinions, introduction of subsidiary questions and removal of systematic bias as well as the application of Delphi in a simulated decision-making context were among the refinements mentioned. Finally, what might be termed a more sophisticated extension of Delphi, Cross-Impact Matrix, was discussed. It was seen that this attempts to discover the conditional probability of an event, given the occurrence or non-occurrence of other events.

Chapter 6

DELPHI IN EDUCATION

PLANNING AND INNOVATION IN HIGHER EDUCATION

The present chapter will deal with the uses of the Delphi Technique in the general area of education, and in higher education in particular. Some of the innovations to which it has contributed will also be reviewed.

Delphi and its variations are applicable to all phases of educational planning, at the federal, state, local, or individual institutional level. For example, a state educational planning office might decide on a building program after using Delphi to consult with the local superintendents, or a university's long-range planning program, which must reconcile the differing views among its departments, could utilize a Delphi process with one or two administrators and a cross-section of departmental representatives as a panel of respondents, and use this method to achieve some reconciliation. On the national level, for another example, the U.S. Office of Education may want to establish a comprehensive program of educational innovations, and this multifaceted problem would provide a number of opportunities for the application of the Delphi Technique (Helmer, 29).

Helmer also stated:

Since educational innovations planned today will probably not be introduced for several years, and since the effects of such innovations--in terms of increased ability among new graduates to cope with the vicissitudes of life--may not be noticed for many years thereafter, decisions regarding such innovations cannot really be made rationally without a reasonably clear image of what the socioeconomic and technological environment of the next few decades

will be.

A study which would place the proper emphasis on those aspects most relevant to education, would provide a useful foundation for planning educational innovations. It seems necessary to establish appropriate educational goals, if educational endeavor is to exert a desired effect on the projected character of the future environment. This appears, largely, to be a matter of preference judgments, that could be obtained through the Delphi Technique, from a cross-section of educators, psychologists, sociologists, and community leaders.

Following such preparatory steps, Helmer continued, a wide survey of suggestions for potential innovations should be made. The resultant compilation would then be pared down by eliminating those items that, on careful consideration, seem to offer little promise of contributing to the previously established educational goals.

A dollar cost estimate of each item on the resulting list of contemplated innovations could then be made. For some of the items, established costing procedures based on current data could furnish reasonably accurate cost estimates. These items would represent simple extensions of present-day practices. For other items, however, the opinions of experts may have to be obtained.

A cost-benefit estimate for individual innovations is considered necessary to determine the overall net benefits expected to be derived from a given dollar expenditure for each proposed innovation. Here again, an attempt to reconcile initially differing opinions could be facilitated by a Delphi inquiry.

Finally, on the basis of these cost-benefit estimates, a program of educational innovations could be constructed by allocating a given budget among

the items on the list of innovative proposals. If these items were entirely independent of one another, an optimal allocation could easily be computed at this point. However, because they are intricately interconnected in terms of costs and benefits, an appeal to judgment must once more be made, and a combined simulation-Delphi use of experts seems to be indicated. (This would be a good place to use Cross-Impact Matrix for interrelationships) (Helmer, 29).

A series of Delphi pilot studies were carried out in an Educational Innovations Seminar held at the Institute of Government and Public Affairs, UCLA, in 1965. Applying the Delphi process, a list of potential educational innovations, together with rough cost estimates for each, was obtained from a total list of 93 distinct proposals. The list was as follows: 1) increase in student participation, 2) educational research and development, 3) model facilities, 4) administration of school systems, 5) internal administration of schools, 6) professional staff, 7) costly new equipment, 8) reorganization of instruction and program, 9) adult retraining, 10) education in the home, and 11) education of deprived populace. The experts were then grouped into several panels and each panel was asked to go through a simulated planning procedure by deciding how a given 5-year budget of 10-billion dollars should be allocated to the educational innovations contained in the list. In order to rationally make these allocations, the participants had to engage in an intuitive cost-benefit appraisal of each item. The best manner in which a group consensus of each such appraisal could be obtained was by the use of a Delphi synthesis of their individual opinions (Helmer, 30; 29).

From these studies, it was concluded that the compilation of a large number of ideas for possible educational innovations served a useful purpose, but not too much weight should be given to the findings from these pilot studies. "Methodologically the endeavor was found very promising by the partici-

pants, who feel encouraged to apply the techniques used to similar problems in a more comprehensive manner in the future" (Helmer, 29).

For a thorough report on the Educational Innovations Seminar, one that is more comprehensive than the report by Helmer (29), the reader is referred to the American Behavioral Scientist, Vol. 10, March, 1967 (Adelson, et al., 1). The discussion is so detailed and extensive that a review would not consider all the variables noted; however, the work flow chart of the seminar is reproduced on the following page, and it may be helpful to the reader in determining whether he wants to learn more about the seminar.

In the absence of an extensive review of the seminar, it might be well to summarize some of the conclusions and observations found.

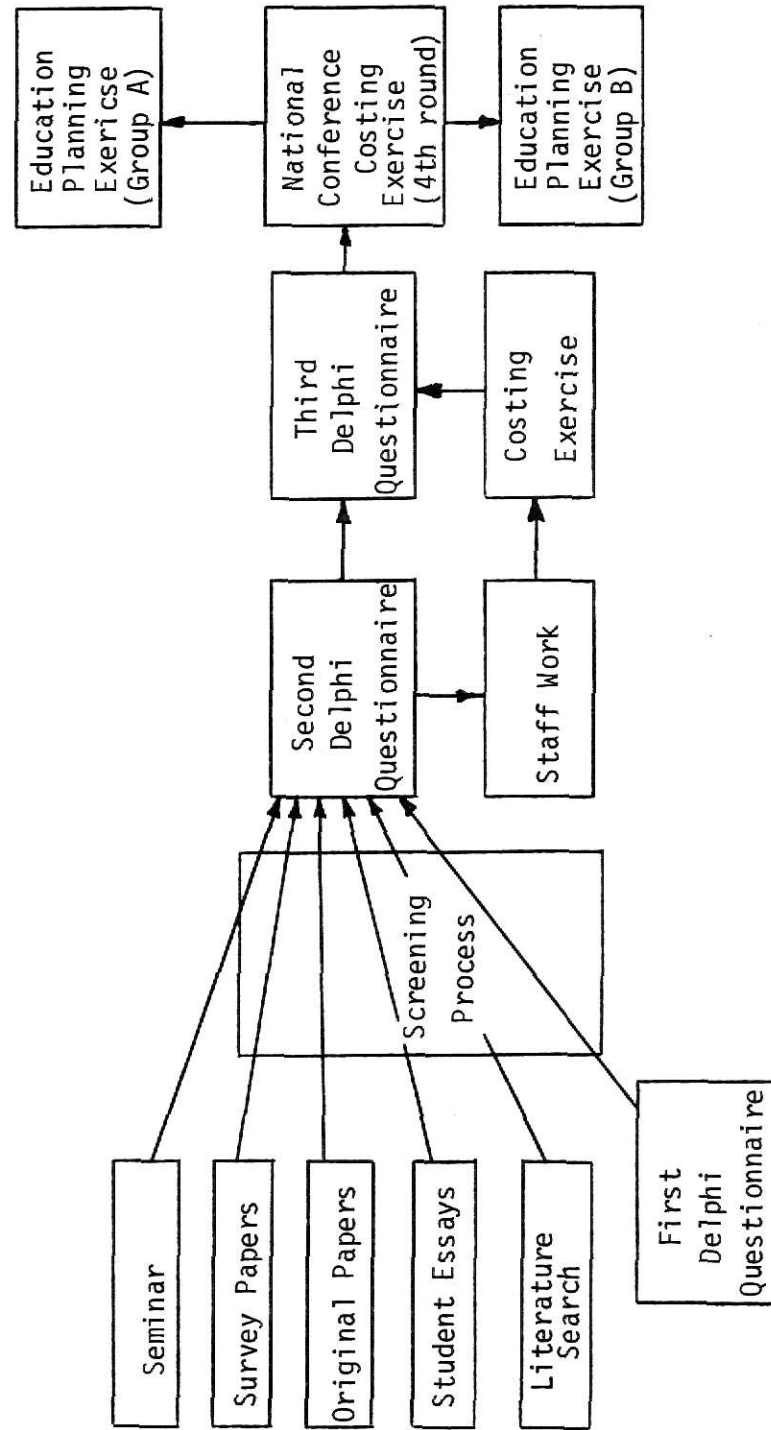
The first and perhaps most sweeping observation was the general agreement among the participants and in the literature on the need for innovation in education. As Adelson, et al., (1) explained:

Very few people argue that American education should be left as it is. Our results seemed to indicate that many innovations are currently believed to be needed. ...In the long run, the need is not merely to innovate, but to develop ways of determining how and when to innovate. ...As a result, we must begin to pay as much attention to institutional arrangements as to technology and curriculum content. Information technologies can have enormous impact in the field of evaluation, for they can help us to know more precisely what it is we are accomplishing and what these accomplishments are costing.

Another finding was that, while the Delphi rounds attempted to produce consensus on preferred alternatives, some lack of consensus was recorded that did not seem to be resolved in the number of iterations used. It is not known whether more rounds of iteration would have resolved these. It was also considered important to improve the decision process in education as much as to modify any of the specific features of contemporary schooling.

In addition, it was felt that it is becoming increasingly feasible,

WORK FLOW CHART FOR EDUCATIONAL INNOVATIONS
SEMINAR



Adelson, et al., "Planning Education For The Future". American Behavioral Scientist,
Vol. 10, No. 7. March 1967, p. 10.

from a technological point of view, to improve the decision-making process within education and its related fields. That is, useful tools and techniques are reaching the status of "shelf items", and the cost-benefit approach is being adapted and its limitations are becoming better understood.

In specific reference to the Delphi Technique, Adelson, et al. (1) stated:

The Delphi Technique is being modified and improved so as to be useful in a variety of ways in educational decision-making. Games and other exercises are being increasingly used, and there are even corporations in the business of supplying them to customers. These techniques should provide vehicles for some of the involvement that is needed to facilitate the adoption of good innovations by making them generally acceptable as they reach the point of usefulness.

The Delphi Technique has also been utilized for collecting opinions in teacher education. In a report by Cyphert and Gant (8), the "need for scientifically assessing the needs, desires, and opinions of clientele was behind the exploration of the potentialities of the Delphi Technique by the School of Education, University of Virginia".

The selection of the sample population for the study was based on a rather arbitrary definition of the power structure of the Commonwealth of Virginia as it related to the school of education in particular. Seven categories were included: 1) faculty of the School of Education and selected graduate and undergraduate student leaders, 2) persons in positions of leadership in the university, 3) off-campus elements, 4) organizational leaders, not necessarily professional educators, 5) influential political figures, senators, representatives, etc., 6) leading newspaper editors and other persons dealing with education, and 8) selected teacher educators in the nation.

The respondents were asked to suggest some prime targets upon which the School of Education should concentrate both its energies and its resources

in the next ten years. The first questionnaire was used to collect a list of items, while the second was to differentiate between the items in hierarchical fashion. Also included with the second questionnaire was a bogus item, to test the hypothesis that "through a process of inserting bogus items and distorting the respondent's reaction to them, it would be possible to mold opinion as well as to collect it". The responses from the second questionnaire were processed by computer to determine the distribution and mode of the priorities assigned to each item. Questionnaire three used this information to report the group consensus and the individual respondent's prior rating for each item. Each person was then asked to re-rate all items in light of this additional information concerning the group feeling. For those items on which the respondent wished to remain outside the consensus, he was asked to state his primary reason for so doing. The ratings on the third questionnaire were computed to note any changes from the second one, and the fourth questionnaire was constructed. It also contained the mode and the individual ratings, plus a synthesized report of the dissenting opinions from questionnaire three. Respondents were then asked to base their final ratings on their own values and a knowledge of both majority and minority views.

Although the authors considered the "mode" as the weaker concept for defining consensus (other studies used the median and inter-quartile range), they felt it was necessitated by the response divergence obtained.

General observations and conclusions from this study regarding the use of Delphi seem worthy of consideration. It was felt that prospective participants must be made to feel that their response is valid if they are to be expected to participate in the study. In addition, the authors commented on the method itself:

There is considerable administrative work and problems as-

sociated with the technique: the maintenance of individual records for each respondent to determine changes and prior ratings; the synthesizing of free responses into communicable generalizations; and the preparation and mailing of several generations of questionnaires and the tabulation of data.

However, the amount of administrative work seems to be directly associated with the size of the group; in this case, there were 400 as opposed to the usual 50.

It was found that the hypothesis about the molding of opinion as well as collecting it was supported by the results obtained. Also, the authors had expected the general movement on the fourth questionnaire to be away from the consensus due to the injecting of minority views. However, this did not materialize, and little need for going beyond the third round was seen.

The authors concluded that the data generated in the study were of value for assisting in formulating the future targets of the School of Education, in addition to making influential persons in the Commonwealth aware of the school's existence and thus awakening them to a realization of its future accomplishments (Cyphert and Gant, 8).

"At present there is a noticeable lack of broad participation in the decision-making and the educational processes of public higher education" (Chalmers, et al., 7). This statement serves to point-up the need for such involvement in higher education.

There are few activities that appear more important to American higher education than systematic and perceptive planning and decision-making. Clear and explicit goals statements need to be provided for the necessary focus and direction of educational institutions (Winstead and Hobson, 51).

"One noticeable trend among researchers looking at institutional goals and objectives...is the use of the Delphi method as a research technique". It has been used, as reported by Winstead and Hobson (51) for determining the

goals or characteristics to be developed for a new university, and to aid in the development of a curriculum for a branch campus of another university.

The National Laboratory for Higher Education, believing that the starting point for effective planning and decision-making should be a clarification and understanding of the goals of an institution, has developed a process for helping institutions of higher education to clarify their goals, gain support for, and derive measurable objectives from, these goals and move toward an administrative style using "management-by-objectives". In an initial study of this process, three administrations of a questionnaire were used and mailed to participating institutions. Each Delphi questionnaire was a list of goal statements comprising eighteen areas. The results of the study showed that there was a remarkable similarity in the institutional profiles obtained, that the instrument was useful as a communication medium among constituent groups, as well as a means of collecting valuable data for the institution's decision-making process, and that the Delphi Technique generally encouraged convergence both within and among the groups (Winstead and Hobson, 51).

However, it was felt that just achieving consensus or reasonable agreement upon goals for diverse groups was just a first step in the planning process. It was suggested that a need existed for institutional goals to be made more explicit and measurable, probably in the form of objectives. As the authors stated,

The process envisioned--that of setting goals, deriving measurable objectives which support these goals, and managing resources to enhance the probability of attaining these goals--will be, in our opinion, a new mode of operation for many educational institutions.

Conclusions regarding the use of Delphi were of the following: The consensus obtained may or may not represent a real shift in attitude, feeling, or perception on the part of certain individuals. For example, it may have

been easier to mark "is" and "should be" on the inventory to correspond with previous modal responses in order not to have to write the reasons for a divergent opinion. In addition, it seems that there is yet no evidence to indicate that the consensus obtained through the process is permanent in nature (Winstead and Hobson, 51).

The final study to be reviewed should be of considerable interest to those readers interested in student personnel and student affairs, in particular, in higher education. The study was concerned with "planning for the future" at Ohio State University and reported on a project involving organizational development in the student affairs department (Rodgers and Lipsetz, 45). The authors explained the concept of "organizational development":

Organizational development is a strategy designed to avoid the pitfalls of a reactive decision-making process. It is a program of planned change that is organization-wide. The key to an organization development project is the use of behavioral science techniques as they relate to the organization. Crucial to the OD plan is that those in the organization with decision making authority be committed to the changes planned. Typically, subordinates prepare proposals for change and submit them to superiors who have no commitment to the proposal and feel that they are free not to implement it.

The authors hoped that the presentation would reveal that the Delphi Technique is useful for the student personnel worker in many facets of his work. It was felt that useful research is scarce for the student personnel worker and that most research in the area more often than not ends, dismally, with "no significant differences". In an organization development model, the organization is of concern in a highly subjective way. "Applying Q-methodology to the data generated by Delphi is a unique approach to data gathering for the student personnel worker. This method allows comparisons to be made on chosen variables, between individuals. It is not concerned with cause and effect relationships."

A task force on the future of student affairs at Ohio State was appointed to consider three missions: 1) to develop a philosophical statement that would guide the future development of student affairs and to subsequently develop new goals, roles and functions for the area, 2) to develop some proposed implementation plans for these goals, roles and functions, and 3) to develop a plan to gain the support of the remainder of the university and others, for the philosophy and programs developed.

After two months of deliberation, meetings and consultations, the task force projected eleven objectives, in an attempt to comply with their mission. One of the more significant objectives was, "to involve all of the staff in processes of the Task Force in order to get better decisions and planning, and decisions that have the personal commitment of the staff".

To implement the task objectives, an action-research model of intervention was used, which involved three processes: 1) the gathering of data from all the individuals and groups within student affairs, 2) feedback to the staff, and 3) joint action planning based upon that feedback.

The Delphi process was selected as the method of data gathering because it permitted all of the staff to participate in defining the future of student affairs and it did so by avoiding face-to-face confrontation (at least initially) and prevented the "pecking order" of the organization from interfering with the presentation of opinion.

The staff of Student Affairs was asked to describe, in the first questionnaire, goals and objectives of an ideal student affairs program at the university. One the second round, each member received a copy of the 28 most frequently mentioned goals and objectives, and was asked to rate or evaluate the items according to the priority he felt they should receive within the organization. The third questionnaire included the list of goals and objectives

as well as the mode rankings for each item, and the staff were requested either to revise their personal opinions and join the mode of the group, or to specify their reasons for remaining outside the consensus. The fourth questionnaire included the list of items, the new mode rankings, and the nature of minority opinions. Each staff member was given a final chance to revise his priorities for the goals and objectives; that is, to join the consensus or to remain outside and re-justify his reasons for so doing.

The final output of this process identified both those goals and functions for which there was agreement and those for which there was disagreement.

The staff of the department then held a three-day retreat to discuss the data obtained from the Delphi procedure. Here, specific groups were brought together to reduce unhealthy competitiveness between the groups and to resolve such inter-group conflicts as overlapping responsibility and confused lines of authority.

During the small group discussions, those areas of consensus were rapidly considered, but those of disagreement consumed much time and energy before they were resolved to the satisfaction of the participants. Upon conclusion of the retreat, a document was produced summarizing the consensus reached by the staff during the retreat.

During the following week, a one-half day planning session was held where priorities were set, commitments to action made, and various planning groups assigned. This resulted in 22 sub-task forces planning the implementation of the goals and objectives produced.

The Delphi process was considered of particular use in implementing the objectives because it involved all of the staff in projecting the future of the organization. Secondly, it provided the staff with definitions of their areas of agreement and conflict, and helped create conditions where this

could be brought out and managed.

The authors, although questioning whether Delphi facilitated real shifts in attitudes, feelings and perceptions on the part of the staff, did feel that the process pinpointed areas of consensus and disagreement, and thus allowed the staff to deal with it in an open, problem-solving climate.

The authors summarized thus, "the Delphi process was a useful intervention technique in the development of the organization of Student Affairs at The Ohio State University" (Rodgers and Lipsetz, 45).

FUTURISTICS

Another indirect, but at times direct use of the Delphi Technique is in the area of the study of the future, called "futuristics". Colleges and universities, as well as high schools and even lower grades have instituted courses in the future; however, it cannot be definitely said whether Delphi actually was responsible for this new orientation, or whether, after introducing a course on the study of the future, some instructor became aware of Delphi and the two have been together ever since.

In the March, 1972 issue of Nation's Schools, futuristics is considered a "crystal ball for curriculum" (20). Studying about the future was seen as about to be regarded as "survival training" in the schools, to aid students of all ages in learning how to cope with the extensive changes they will encounter in the future.

The Delphi method, along with scenarios, simulations, games, and trend extrapolations were seen as useful tools for teachers of the future. Delphi is considered simple to use and easily adapted to the classroom situation. At the University of Massachusetts, the method has been used in classes, and is reported to generate discussions almost immediately, mainly because in predicting

the future, "there are no right answers". As one person, quoted in the article, stated: "Whatever its shortcomings as a forecasting tool, the Delphi is a good introductory technique because it gets people talking about the future and sharing their perceptions".

A "scenario" is a narrative that combines the future with a history leading up to it. It has been used mostly by professional futurists to design alternative futures, and can serve the same purpose in the futuristics class. The future state of some system, for example, is described, and a probable history of the evolution of the current state of that system to its future state is drawn or outlined.

Trend extrapolation involves identification of current trends in a given area which are then projected into the future. Students can also construct probable futures by extrapolating several trends, and relying on their own knowledge and appropriate data. A desirable future can then be compared to this probable future, and a scenario approach can be used to discuss what needs to be done in the present to shift these probable futures to desirable futures.

Another reference to "futuristics" was found in the Spring, 1971 issue of Trend, all of which was devoted to the subject (Rojas, 46). In this article, Alvin Toffler, in an interview, discussed the future of education and noted some trends affecting education in the upcoming years. He saw the introduction of the future into education as a way of "healing some of the deep wounds of the future" as well as a way of "healing some of the deep wounds in education today".

The introduction of the future, according to Toffler, can open doors to general educational reform, as well as allowing schools and colleges to use futurism as a lever for institutional and curricular reform. On the topic of

values, Toffler felt that it was important for students to explicitly state their own values, and any educational process that aids in this can help them to clarify their values and also help them in their personal relations as well as in decision-making.

Another section of the ~~J~~Journal concerned descriptions of some current educational programs involving Futuristics. Suffice it to say that well over one hundred courses focusing on the future are currently being offered in the United States and Canada, and educational interest in Futuristics seems to be developing throughout the country. Indeed, there are a number of books that can be used as texts for future courses and there is even a game, designed by Olaf Helmer and Theodore Gordon (who else?), called FUTURES, which was once distributed by Kaiser Aluminum. Unfortunately, it is no longer available.

In discussing the study of the future in education, one is tempted to try his own hand at some predictions concerning the future of education itself. However, that task is spared us because of a number of comments in this regard that were encountered in the literature on Delphi. A few of these seem worthy of promotion and will be offered, out of context, for the reader's consideration.

Hayden (25) made the following comments in regard to the future of education:

We can expect enormous changes in education--not only in techniques, but in its impact on our whole society. ...[Education will be an on-going thing]...universities already have what they call "centers for Continuing Education" and so-called "extension" courses reach out into the adult community.

The proportion of the population that will have had a college education will pass the two-thirds mark in 1980, and continue to rise. Colleges will become a place for retraining, as adults, trying to keep pace with advancing technology, are forced to learn to do 2 or 3 or more kinds of jobs during their working career.

And, Helmer (32) saw it from a different perspective:

There are signs today of widespread recognition that the complex problems involved in shaping the future of our society are multidisciplinary in character and that their solution requires the collaboration of scientists and technologists from many different fields.

The universities are unable to spearhead...[the movement]...toward true interdisciplinary cooperation, largely because of the incentive structure that discourages scholars from devoting too much time to enterprises that do not find their rewards within the departmentalized promotion system. ...[However]...there are strong indications that the universities will undergo the necessary administrative and curriculum reforms to enable them ...[to take over the reins].

SUMMARY

The present chapter was concerned with the application of the Delphi Technique in the area of education. Several studies of its uses in higher education were reviewed, and it was found that Delphi was useful for defining institutional goals and as an aid in decision-making and planning.

It was also shown that Delphi seemed to have had, at least, an indirect impact on the study of the future, called Futuristics, that has recently acquired interest by colleges, universities, high schools, and even grade schools.

Finally, a few speculative comments concerning the future of education were offered for whatever their inherent worth.

Chapter 7

EVALUATION

This final chapter will be concerned, mainly, with some general comments regarding a few aspects of the Delphi Technique, and a review of a critique of the method found in the literature. There will also be a few comments regarding some suggested improvements in the technique for further research.

STRUCTURED FACE-TO-FACE INTERACTION AND ANONYMOUS QUESTIONNAIRES

Dalkey and Rourke (33) reported an initial experiment designed to compare the efficacy of the Delphi Technique with face-to-face discussion. It was indicated that the first-round, "off-the-cuff" estimates of the subjects were at least as accurate as the consensus reached by the face-to-face group after half-an-hour of discussion.

Dalkey (13) reported that there was a "much richer" exchange of information during the face-to-face encounter than in the Delphi sessions, but that degradations due to the effect of dominant individuals, semantic noise, and group pressure were evident.

Dalkey (14) concluded that more often than not, the Delphi procedures improved the group response compared to those obtained with face-to-face discussion.

VALIDITY

*need more
to defend
validity*

The definition of validity in the context of the Delphi Technique is concerned with the method yielding more reliable results than rival, more traditional, methods. One reason given is that, as Delphi was initially intended for use with experts, it is difficult to persuade actual experts to participate in an experiment. It is not known, furthermore, whether favorable results, obtained through the use of students as "experts" are transferable to the real item. "Therefore, the validity of the procedure may be considered established only in an intuitive sense, in that the participants themselves... generally appear satisfied that the method is both fair and efficient in extracting...information..." (Dalkey and Brown, 35).

RELIABILITY

If one accepts the notion of reliability as referring to the range of estimates obtained, then there seems to be mounting evidence that the systematic processing of expert opinion produces significant improvement in both accuracy and reliability. However, there appear to be no cases where the relative effectiveness of the Delphi Technique has been compared to other traditional forms (Dalkey and Rourke, 33).

The concept of reliability can be better discussed within the context of the n-heads rule, discussed in Chapter 1 (p. 5). Specifically, from the standpoint of the decision-maker, the most uncomfortable aspect of opinion appears to be that experts with apparently equivalent credentials (equal degrees of expertness) are likely to give quite different answers to the same question. One of the major advantages of using a group response is that this diversity of answers is replaced by a single representative opinion (Dalkey,

15).

Dalkey (15) stated it another way:

In general, one would expect that in the area of opinion group responses would be more reliable than individual opinions, in the simple sense that two groups (of equally competent experts) would be more likely to evidence similar answers to a set of related questions than would two individuals.

Dalkey further contended that the similarity in answers could be measured by a correlation between the answers of the two groups in regard to a set of questions. However, it is not a tautological assertion that groups will be more reliable than individuals. This would depend on the distribution of answers obtained, and upon the method of selecting the sub-groups from the total population. The assertion could be expected to be supported if the distributions of answers are not highly distorted and if the sub-groups are selected at random.

Dalkey (15) concluded that,

For the analyst using expert opinion within a study, reliability can be considered to play somewhat the same role as reproducibility in experimental investigations, ...[since]...it is clearly desirable for a study that another analyst using the same approach (and different experts) arrive at similar results.

In one study, Dalkey continued, it was clear that there was a definite and monotonic increase in the reliability of the group responses when the size of the group increased. This could infer that the more "experts" one has, the more reliable will be the eventual outcome of a Delphi exercise.

CONSISTENCY

An article, purporting to investigate the consistency of Delphi forecasts was found in the April, 1970 issue of The Futurist (Martino, 39), by the technological forecasting editor of that magazine.

Martino saw one of the disadvantages of Delphi as the length of time

Panelist dropout is not a well known hazard. Although he cited no empirical data, Martino (1972) asserted that response rates to 1st round question "typically run 50% or less".

74

required to obtain several iterations from a panel, but concentrated his comments on the consistency of the forecasts produced by the technique.

Martino did not consider the accuracy of a forecast as important, but rather the utility of the forecast as information to be used for decision-making. In considering the possibility of two different groups of experts at two different times producing divergent forecasts, he questioned the likelihood of that occurrence as a measure of the consistency of those forecasts.

Upon comparison of several Delphi studies, Martino concluded that in general, Delphi forecasts seem to be consistent. That is, "different panels of experts on a subject do tend to produce about the same forecasts", and he considered it unlikely that "two panels of equally competent experts will produce forecasts which are significantly different" (Martino, 1972, 39).

A rather extensive and not-at-all flattering critique of the Delphi Technique, at least as it is used in forecasting, was presented in the January, 1971 issue of the Phi Delta Kappan (Weaver, 50).

Weaver contended that the use of forecasting must be incorporated in educational planning and policy decisions where it would be necessary to continually conduct and assess studies of the future. The main emphasis was on the grounds upon which educators can select among the tools for forecasting.

Weaver explained:

At a minimum, the value of such methodologies cannot be weighed apart from some reasonable understanding of the processes by which they aid in producing forecasts. ...In order that the results of the forecasts can be effectively evaluated, one must test the plausibility of the forecasts, which demands that the forecast include some explanatory quality.

To do so, one must look at the reasonableness of the arguments and explanations which support a forecast within the context of forecasting methodologies.

Weaver contended that the Delphi Technique can be effectively critiqued if one considers both its process and its product, that is, by the way people are asked to handle information concerning the future (process), and by a concern with plausibility (product).

Weaver considered two common properties shared by the various methods of forecasting, of which Delphi is only one: 1) the methods are all non-data based, and 2) they rely on collective expert judgment. It is considered crucial that Delphi, along with the other methods, heavily emphasize the explanations upon which the plausibility of the forecasts rest. The singular weakness of Delphi forecasts is seen as due to their lack of "substantive explanatory quality".

In elaborating, it was contended that forecasts have been considered plausible only on the basis of expert consensus or agreement, which is not sufficient for arguing the plausibility or convincibility of the forecasts.

Concerning educational policy, Weaver saw the meaning of a forecast as manifested in what the statement "convinces us the future can be made to be through reasonable actions, not in what it will be". It seems fundamental that unless forecasts "open options", they shall be of little value to policy makers.

Weaver made explicit a disadvantage of Delphi:

The very basis of the Delphi forecasting process is opinion as to when an event is likely to occur. ...[In establishing plausibility of forecasts, it seems essential that]...opinion be distinguishable as rational judgment rather than guess work. Delphi, at present, can render no such distinction, because the arguments which support an opinion are not emphasized unless the opinion is contrary to the group norm.

However, Weaver did end on a positive note, at least for those concerned with education, in that he saw the more promising uses of Delphi in educational applications as, "(a) a method for studying the process of think-

ing about the future, (b) a pedagogical tool or teaching tool which forces people to think about the future in a more complex way than they ordinarily would, and (c) a planning tool which may aid in probing priorities held by members and constituencies of an organization" (Weaver, 50).

SUGGESTED IMPROVEMENTS FOR FURTHER RESEARCH

Gordon and Helmer (23) outlined some suggestions for improvements in the Delphi Technique that might serve as bases for further research.

One area would be experimentation with various methods of information feedback which would aid in learning more about the sensitivity of changes in opinion due to the form and content of such feedback.

Another of the areas worthy of experimentation is the comparative analysis of social pressure and persuasive reasoning as they are related to the convergence of opinion. A statistical model could be formulated of the question-and-answer operation, as another area of study, in which the operation could be viewed as a measuring instrument "for the substantive quantities which form the subject of the questions; each respondent would here have to be represented by an error distribution, and some hypotheses would have to be stated as to the relative independence of the measurement thus obtained".

A final area of exploration, as suggested by Gordon and Helmer, is the development of techniques "for the formulation of sequential questions that would probe more systematically into the underlying reasons for the respondents' opinions, in a deliberate effort to construct a theoretical foundation for the phenomena under inquiry".

In concluding this review of the literature on the Delphi Technique, a statement by Dalkey (15) seems very appropriate. He felt that,

...all of [the] features of a Delphi exercise are desirable, especially if the exercise is conducted in [a] context... where group acceptance is an important consideration. Like any technique for group interaction, the Delphi procedures are open to various misuses; much depends on the standards of the individual or group conducting the exercises.

REFERENCES

1. Adelson, M., et al. "Planning Education for the Future," American Behavioral Scientist, 10:1-29, March, 1967.
2. Asch, S.E., "Effects of Group Pressure Upon the Modification and Distortion of Judgments," eds. E.E. Maccoby, et al. Readings in Social Psychology. 3d ed. London: Holt, Rinehardt, and Winston, 1958.
3. Brown, B., and O. Helmer. Improving the Reliability of Estimates Obtained From a Consensus of Experts. P-2986. Santa Monica, California: The RAND Corporation, September, 1964.
4. Brown, B. DELPHI PROCESS: A Methodology Used for the Elicitation of Opinions of Experts. P-3925. Santa Monica, California: The RAND Corporation, September, 1968.
5. Cambell, R.M., and D. Hitchin. "The Delphi Technique: Implementation in the Corporate Environment," Management Services, 5:37-42, November-December, 1968.
6. Cantril, H. "The Prediction of Social Events," Journal of Abnormal Psychology, 33:364-389, July, 1938.
7. Chalmers, J., et al. "Societal Developments: Implications for State Universities and Land-Grant Colleges" Unpublished preliminary report for the National Association of State Universities and Land-Grant Colleges, Commission on Arts and Sciences, September, 1970.
8. Cyphert, F.R. and W.L. Gant. "The Delphi Technique: A Tool for Collecting Opinions in Teacher Education," Journal of Teacher Education, 21:417-425, Fall, 1970.
9. _____. "The Delphi Technique: A Case Study," Phi Delta Kappan, 52:267-271, January, 1971.
10. Dalkey, N., and O. Helmer. "An Experimental Application of the Delphi Method to the Use of Experts," Management Science, 9:458-467, April, 1963.
11. Dalkey, N. Delphi. P-3704. Santa Monica, California: The RAND Corporation, October, 1967.
12. _____. Quality of Life. P-3805. Santa Monica, California: The RAND Corporation, March, 1968.

- ✓ 13. _____. Experiments in Group Prediction. P-3820. Santa Monica, California: The RAND Corporation, March, 1968.
- ✓ 14. _____. Predicting the Future. P-3948. Santa Monica, California: The RAND Corporation, October, 1968.
- ✓ 15. _____. The Delphi Method: An Experimental Study of Group Opinion. RM-5888-PR. Santa Monica, California: The RAND Corporation, June, 1969.
- ✓ 16. _____, et al. The Delphi Method, III: Use of Self-Ratings to Improve Group Estimates. RM-6115-PR. Santa Monica, California: The RAND Corporation, November, 1969.
- ✓ 17. _____, et al. Measurement and Analysis of the Quality of Life: With Exploratory Illustrations of Applications to Career and Transportation Choices. RM-6228-DOT. Santa Monica, California: The RAND Corporation, August, 1970.
- ✓ 18. _____, and D.L. Rourke. Experimental Assessment of Delphi Procedures with Group Value Judgments. R-612-ARPA. Santa Monica, California: The RAND Corporation, February, 1971.
19. _____, and B. Brown. Comparison of Group Judgment Techniques with Short-Range Predictions and Almanac Questions. R-678-ARPA. Santa Monica, California: The RAND Corporation, May, 1971.
20. "Futuristics: Crystal Ball for Curriculum," Nation's Schools, 89:3, March, 1972.
21. Glushkov, V.M. "Forecasting Events on the Basis of Expert Opinions," Soviet Cybernetics Review, ed. W.B. Holland. RM-6000/10-PR. Santa Monica, California: The RAND Corporation, October, 1969.
22. Goldsen, J.M. Thinking About Future Social Development. P-3680. Santa Monica, California: The RAND Corporation, September, 1967.
23. Gordon, T. and O. Helmer. Report on a Long-Range Forecasting Study. P-2982. Santa Monica, California: The RAND Corporation, September, 1964.
24. _____, and H. Hayward. "Initial Experiments with the Cross-Impact Matrix Method of Forecasting," Futures, 1:100-116, December, 1968.
25. Hayden, B. The Year 2000. P-3571. Santa Monica, California: The RAND Corporation, March, 1967.
- ✓ 26. Helmer, O., and N. Rescher. "On the Epistemology of the Inexact Sciences," Management Science, 6:25-52, October, 1959.
27. _____. Convergence of Expert Consensus Through Feedback. P-2973. Santa Monica, California: The RAND Corporation, September, 1964.
28. _____. A Use of Simulation for the Study of Future Values. P-3443. Santa Monica, California: The RAND Corporation, September, 1966.

29. _____. The Use of the Delphi Technique in Problems of Educational Innovations. P-3499. Santa Monica, California: The RAND Corporation, December, 1966.
30. _____. Analysis of the Future: The Delphi Method. P-3558. Santa Monica, California: The RAND Corporation, March, 1967.
31. _____. New Developments in Early Forecasting of Public Problems: A New Intellectual Climate. P-3576. Santa Monica, California: The RAND Corporation, April, 1967.
32. _____. The Future of Science. P-3607. Santa Monica, California: The RAND Corporation, May, 1967.
33. _____. Methodology of Societal Studies. P-3611. Santa Monica, California: The RAND Corporation, June, 1967.
34. _____. Prospects of Technological Progress. P-3643. Santa Monica, California: The RAND Corporation, August, 1967.
35. _____. Systematic Use of Expert Opinions. P-3721. Santa Monica, California: The RAND Corporation, November, 1967.
36. Kaplan, A., et al. "The Prediction of Social and Technological Events," Public Opinion Quarterly, 14:93-110, Spring, 1950.
37. Kelley, H., and J. Thibaut. "Experimental Studies of Group Problem Solving and Process," ed. G. Lindzey. The Handbook of Social Psychology, Vol. II. Cambridge, Mass.: Addison-Wesley Pub. Co., Inc., 1954.
38. Maier, N. "Assets and Liabilities in Group Problem Solving: The Need for an Integrative Function," Psychological Review, 74:239-249, July, 1967.
39. Martino, J. "The Consistency of Delphi Forecasts," The Futurist, 4:63-64, April, 1970.
40. McGregor, D. "The Major Determinants of the Prediction of Social Events," Journal of Abnormal Psychology, 33:179-204, January, 1938.
41. Quade, E. An Extended Concept of "Model". P-4427. Santa Monica, California: The RAND Corporation, July, 1970.
42. Rescher, N. Delphi and Values. P-4182. Santa Monica, California: The RAND Corporation, September, 1969.
43. _____. The Future as an Object of Research. P-3593. Santa Monica, California: The RAND Corporation, April, 1967.
44. Rochberg, R. Some Comments on the Problem of Self-Affecting Predictions. P-3680. Santa Monica, California: The RAND Corporation, September, 1967.

45. Rodgers, R., and A. Lipsetz. "Planning for the Future: An Organization Development Project in the Student Affairs Department at the Ohio State University" Unpublished paper. Columbus, Ohio: Ohio State University, September, 1967.
46. Rojas, B. (ed.). "Futuristics in Education," Trend, 7:3-26, Spring, 1971.
47. Simon, H. "Bandwagon and Underdog Effects," Models of Man. New York: John Wiley and Sons, Inc., 1957.
48. Turoff, M. "The Delphi Conference," The Futurist, 5:55-57, April, 1971.
49. Umpleby, S., and J. Briggs. "Exploring the Future with a Computer," The Futurist, 4:197-199, December, 1970.
50. Weaver, W. "The Delphi Forecasting Method," Phi Delta Kappan, 52:267-271, January, 1971.
51. Winstead, P., and E. Hobson. "Institutional Goals: Where to from Here?," Journal of Higher Education, 42:669-671, November, 1971.
52. Wishart, B. "Delphi Technique: Integral Project Resource Element," Thrust, 1:45-46, December-January, 1972.

THE DELPHI TECHNIQUE: A REVIEW OF
THE LITERATURE

by

ARTHUR JOSEPH PELLETIER

B. A., Kansas State University, 1970

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

College of Education

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1972

ABSTRACT

Institutions of higher education are dependent from time-to-time on the decisions handed down by various committees, or commissions, or other such groups. One obvious characteristic of these groups has been that they are usually in the form of a face-to-face encounter. A number of investigators have found, however, that there are a number of disadvantages to such interaction, in the form of certain social-psychological phenomena such as the effect of the dominant individual, or the unwillingness of some members of the group to abandon expressed opinions, or the effect of redundant or irrelevant material called "semantic noise".

The Delphi Technique, developed at the RAND Corporation, is a device for overcoming some of these disadvantages by utilizing a series of prepared, sequential, anonymous, questionnaires, alternately interspersed with opinion feedback and other information. The technique was originally designed as a forecasting tool to be used with experts for the prediction of the future. Since its emergence, however, it has been utilized in a number of ways and a variety of experiments have been conducted to investigate not only the characteristics of the technique, but its applications to the study of the future, of values, and of the quality of life, to mention a few. It has also been used to define institutional goals, and organizational objectives.

As with any new technique, Delphi is not without its disadvantages. The more notable of these is the amount of time required to prepare and mail several generations of questionnaires and to compile the data from each. How-

ever, the use of the technique in conjunction with computers has eliminated this problem.

The present paper was a review of the literature concerning the Delphi Technique. The development of this device from initial studies on the prediction of social events, to its intended use as a forecasting tool, to latter-day refinements, was traced. A number of experiments in the use of the technique were reviewed, although only a few were described beyond general statements of purpose and results obtained. Since many were extensive and detailed in nature, an in-depth review would burden the paper with unnecessary repetition. Also, in the interest of easy reading, myriads of statistics, graphs, charts, and the like were excluded.

An extensive list of references was provided, which together with the references provided by the articles listed can assure the reader of having sufficient data to draw his own conclusions concerning the Delphi Technique.