ANALYSIS OF THE FUTURE: THE DELPHI METHOD

Olaf Helmer

March 1967

P-3558

ANALYSIS OF THE FUTURE: THE DELPHI METHOD

Olaf Helmer*

The RAND Corporation, Santa Monica, California

The decade of the Sixties has brought with it an important change in the intellectual climate throughout many parts of the world, evidenced by a new attitude toward the future that has become apparent in public and private planning agencies as well as in the research community. The effect 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.

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.

Let me expand a little regarding these three aspects of our thinking about the future.

---

*Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The RAND Corporation as a courtesy to members of its staff.

This paper was prepared to be published this year in Italian in <u>Rivista Italiana di Amministrazione Industriale</u>.

By the change in philosophical attitude to which I referred I mean that the exploration of the future is no longer equated with fortune-telling or with crystal-ball gazing. Instead there is a growing awareness that a great deal can be said about future trends in terms of probability, and moreover that through proper planning we can exert considerable influence over these probabilities. Fatalism, in other words, has become a fatality. 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.

As for the new pragmatic attitude to which I had also referred, and which is beginning to be noticeable in government as well as in industry, it is due--I think--to the fact that not only are technology and our environment undergoing change but the pace of change in our time is accelerating. No longer does it take generations for a new pattern of living conditions to evolve, but we are going through several major adjustments in our lives, and our children will have to adopt continual adaptation as a way of life. For such adaptation to occur without major psychological or economic disruption, it is becoming mandatory for us to strive to anticipate changes in our environment rather than to attempt to deal with them belatedly and inadequately after it has become obvious that they are upon us. The recognition of this need for anticipation has had visible effects. Until not so long ago, systematic efforts at long-range governmental or industrial planning were in bad odor in the capitalist countries, because they were associated in the public mind with what many considered the worst aspects of state-controlled socialism. This view, fortunately, is now a matter of the past, and there is a general awareness that in the competition between Western capitalism and the Communist bloc the former cannot hope to prevail unless the quality of its long-range planning is unsurpassed. As evidence I can point to several developments, in labor, in industry, and in government. Labor unions show growing explicit concern not just over the near future but over the long-term social implications of automation and other causes of increased productivity. In industry there is likewise a noticeable change in attitude, inasmuch as a hard look is beginning to be taken in many places at the possible long-term futures of our society, in order to derive from such forecasts appropriate guidance for the operation of industrial corporations. As for government, finally, the various Departments of the federal government in the United States are undergoing a thorough reorganization of their planning methods; this is in consequence of a directive from President Johnson to introduce throughout so-called program-budgeting procedures first used successfully only in the Defense Department.

The third point I mentioned, our growing ability to do something about the future, I would like to discuss in a little more detail.

In my opinion the so-called soft sciences are on the verge of a revolution. The traditional methods of the social sciences are proving inadequate to the task of dealing effectively with the ever-growing complexity of forecasting the consequences of alternative policies and thus furnishing useful planning aid to high-level decision-makers in the public and private sectors. This situation is now rapidly being remedied, by introducing new methods developed elsewhere in the form of operations research techniques, such as the construction of mathematical models, simulation procedures, and a systematic approach to the utilization of expert opinions, - the latter a subject on which more will be said below. In addition to these techniques, new uses of computers, with automated access to central data banks, will provide the soft sciences with the same kind of massive data-processing capability that, in the physical sciences, created the breakthrough which led to the development of the atomic bomb.

Among the new methods mentioned above that are under development is one that has become known as the Delphi Technique, which attempts to make effective use of informed intuitive judgment. It derives its importance from the realization that projections into the future, on which public policy decisions must rely, are largely based on the personal expectations of individuals rather than on predictions derived from a well-established theory. Even when we have a formal mathematical model available—as is the case, for example, for various aspects of the national economy—the input assumptions, the range of applicability of the model, and the interpretation of the output all are subject to intuitive intervention by an individual who can bring the appropriate expertise to bear on the application of the model. In view of the absence of a proper theoretical foundation and the consequent inevitability of having, to some extent, to rely on intuitive expertise—a situation which is still further compounded by its multidisciplinary characteristics—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 socioeconomic 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.

The best we can do, under the circumstances, when we do have to rely on expert judgment, is to make the most constructive and systematic use of such opinions. In dealing with experts, there are basically three rules which I think ought to be followed: (1) Select your experts wisely. (2) Create the proper conditions under which they can perform most ably. (3) If you have several experts on a particular issue available, use considerable caution in deriving from their various opinions a single combined position. The so-called Delphi technique, explained later, deals with this last point. But first, I will comment on the other two.

It is obvious that much depends on how expert the experts are—their proper selection presents many problems. I will not go into these now, but merely point out that there are difficulties in defining qualifications and measuring relative performance of experts. That is, it is far from obvious what we mean—or should mean—when we say that somebody is an expert; and even given reasonable criteria of expertness it may not be easy to obtain adequate data for determining a person's degree of expertise.

The second rule, that an expert should be placed in the right conditions in order to perform well, means that communication should be facilitated as much as possible. Here, first of all, the prior formulation of an appropriate model (even a very tentative one of the operations-analytical kind) would serve to communicate the problem to him with clarity and receive his answer without risk of misinterpretation. Secondly, the expert would be greatly aided in his performance if he had ready access to relevant information that may exist elsewhere (in this regard, rapid progress in data processing may open up new possibilities by which the present swamping w .h irrelevancies will eventually be replaced with push-button availability of pertinent data in the form of automated libraries). Thirdly, 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.

A particularly effective way of encouraging interaction among experts is to place them in a laboratory situation where they are required to participate in a simulation exercise. In a simulation model a kind of conceptual transference takes place. Instead of describing a situation directly, each of its elements is simulated by substituting a mathematical or physical object for the real one and simulative relations for those that really exist. For example, a policy planning operation can be simulated by a set of make-believe decision makers who, playing roles in a laboratory "game," might go through the decision-making motions that their real-life counterparts would be expected to carry out in actuality.

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. This interjection of a model has the advantage that it admits of what may be called "pseudo-experimentation" ("pseudo" because the experiments are carried out in the model, not in reality).

Past experience with simulation models suggests that they can be highly instrumental in motivating participants to communicate effectively with one another, to learn more about the subject matter by viewing it through the eyes of persons with backgrounds and skills different from their own, and above all to acquire an integrated overview of the problem area. This stimulating effect of collaborating on the employment of a simulation model is particularly powerful when the simulation takes the form of an operational game where the participants act out the roles of decision- and policy-making entities. By being exposed within a simulated environment to a conflict situation involving an intelligent opposition, the "player," no matter how narrow his specialty, is compelled to consider many aspects of the scene that might not normally influence his opinions to the same extent as they do when he works in isolation.

After this excursion into the question of how best to provide an expert with a suitable environment in which to function, let me return to the third rule for dealing with experts: this has to do with the problem of combining the opinions of the members of a panel of experts into a single position.

Perhaps the traditional and in many ways the simplest method of achieving a consensus has been to conduct a round-table discussion among the experts and have them arrive at an agreed-upon group position. This procedure is open to a number of objections. In particular, the outcome is apt to be a compromise between divergent views, arrived at all too often under the undue influence of certain psychological factors, such as specious persuasion by the member with the greatest supposed authority or even merely the loudest voice, the unwillingness to abandon publicly expressed opinions, and the bandwagon effect of majority opinion.

In recent years we have been experimenting with a new approach to overcome these difficulties, which has become known as the Delphi technique. The Delphi technique, in its simplest form, eliminates committee activity among the experts altogether and replaces it with a carefully designed program of sequential individual interrogations (usually best conducted by questionnaires) interspersed with information and opinion feedback.

It may perhaps be easier to describe the principles involved in this procedure by reference to a particular example. When inquiring into the future of automation,* each member of a panel of experts in this field was asked to estimate the year when a machine would become available that would comprehend standard IQ tests and score above 150 (where "comprehend" was interpreted behavioristically as the ability to respond to printed questions possibly accompanied by diagrams). The initial responses consisted

---

*As part of a long-range forecasting study conducted with the participation of Theodore Gordon under the auspices of The RAND Corporation; a report on this study appeared as in a set of estimates spread over a sizeable time-interval, from 1975 to 2100. A follow-up questionnaire fed back to the respondents a summary of the distribution of these responses by stating the median and—as an indication of the spread of opinions—the interquartile range (that is, the interval containing the middle 50% of the responses). The respondent was then asked to reconsider his previous answer and revise it if he desired. If his new response lay outside the interquartile range he was asked to state his reason for thinking that the answer should be that much lower, or that much higher, than the majority judgment of the group.

Placing the onus of justifying relatively extreme responses on the respondents had the effect of causing those without strong convictions to move their estimates closer to the median, while those who felt they had a good argument for a "deviationist" opinion tended to retain their original estimate and defend it.

In the next round, responses (now spread over a smaller interval) were again summarized, and the respondents were given a concise summary of reasons presented in support of extreme positions. They were then asked to revise their second-round responses, taking the proffered reasons into consideration and giving them whatever weight they thought was justified. A respondent whose answer still remained outside the interquartile range was required to state why he was unpersuaded by the opposing argument. In a fourth, and final round these criticisms of the reasons previously offered were resubmitted to the respondents, and they were given a last chance to revise their estimates. The median of these final responses could then be taken as representing the nearest thing to a group consensus. In the case of the high-IQ machine, this median turned out to be the year 1990, with a final interquartile range from 1985 to 2000. The prodedure thus caused the median to move to a much earlier date and the interquartile range to shrink considerably,

---

an appendix to "Social Technology" by O. Helmer (Basic Books, 1966).

presumably influenced by convincing arguments.

This convergence of opinions has been observed in the majority of cases where the Delphi approach has been used. In a few of the cases where no convergence toward a relatively narrow interval of values took place, opinions began to polarize around two distinct values, so that two schools of thought regarding a particular issue seemed to emerge; this may have been an indication that opinions were based on different sets of data or on different interpretations of the same data. In such cases, it is conceivable that a continuation of the Delphi process through several more rounds of anonymous debate-by-questionnaire might eventually have tracked down and eliminated the basic cause of disagreement and thus led to a true consensus. But even if this did not happen, or if the process were terminated before it had a chance to happen, it should be realized that the Delphi technique would have served the purpose of crystallizing the reasoning process that might lead to one or several positions on an issue and thus help to clarify the issue even in the absence of a group consensus.

The illustration given above is intended to describe the basic essentials of the Delphi technique. Refinements are made to fit each particular case; two of them are discussed below.

One is that of introducing weighted opinions. If it were easy to measure the relative trustworthiness of different experts objectively, we would obviously give greatest, if not exclusive, weight to the opinions of those who are most trustworthy. In view of the absence of such measurements, experiments have been carried out to test the degree of reliance that may be placed on the experts' self-appraisal of their relative competence. We found the results to be quite promising. This device was used in November, 1965, when twenty members of the faculty of the Graduate School of Business Administration at the University of California

(Los Angeles) made forecasts of ten economic and business indices for the last quarter of 1965 and for the entire year 1966 (twenty answers altogether). The procedure was as follows: In addition to going through four rounds of Delphi arguments, the respondents were asked to rank their relative competence with regard to the estimation to each of the ten indices. Then, instead of using for each index the median of all twenty final responses as the group consensus, and thus as the group's prediction for 1966, we took only the responses of those individuals who had ranked themselves relatively most highly competent for that particular index, and then used the median of just these forecasts as the group consensus. It subsequently turned out that this select median, compared to the median of all responses, was closer to the true value in 13½ out of the 20 cases.

Secondly, and finally, let me point out a slightly more sophisticated use of the Delphi approach, where it is used in conjunction with a simulated decision-making process of the kind mentioned earlier. A typical situation to which this mode of using expertise is applicable is one in which budgetary decisions have to be made on the basis of cost-benefit estimates.

When costs and benefits are clearly measurable objective terms there is no need to resort to the use of mere opinions. But in practice, benefits resulting from the choice of given policy alternatives are almost never capable of unambiguous measurement; even in the case of cost estimates it is usually only the dollar expenditures which are closely predictable, while social costs may be as elusive as the benefits. In such cases, a consensus of judgments made by experts may be helpful in obtaining an appraisal.

In a recent experiment conducted in the course of a project concerned with educational innovations, expert opinions were used in a context of this sort. Applying a Delphi process, a list of potential educational innovations, together with rough cost estimates for each, was first obtained. We grouped our experts into several panels and asked each panel to go through a simulated planning process by deciding how a given budget should be allocated to the educational innovations contained in the given list. In order to make these allocations rationally, the participants had to engage in an intuitive cost-benefit appraisal of each item on the list. The manner in which a group consensus of each such appraisal could best be obtained was by way of a Delphi synthesis of their individual opinions.

These examples are intended merely to illustrate the potentialities of the Delphi technique. Numerous further experiments need to be carried out 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, in particular, for policy applications in the general area of social technology.