WORKING PAPER

WORKING PAPER

HOW TO DESIGN, STRUCTURE

AND PERFORM A DELPHI STUDY

DOCUMENT DE TRAVAIL

Industry, Trade and Commerce
Industrie et Commerce

Office of Science and Technology
Ottawa, Canada

Direction des sciences et de la technologie
Ottawa, Canada

WORKING PAPER

HOW TO DESIGN, STRUCTURE AND PERFORM A DELPHI STUDY

FRANK J. DOYLE

Office of Science and Technology,
Department of Industry, Trade
and Commerce.
July 1974.

THE DELPHI TECHNIQUE

Delphi has become one of the most widely used forecasting methods and is discussed at some length in a number of technological forecasting texts such as Jantsch in "Technological Forecasting in Perspective", Bright in "Technological Forecasting for Industry and Government" and most thoroughly by Martino in "Technological Forecasting for Decision Making".

The dissertations on Delphi by these authors concentrate on outlining the background, validity and merit of the technique but do not explicitly describe in detail how to design, structure, write and administer a Delphi study. The goal of this paper is to provide pragmatic guidance, based upon the experience gained in eight major studies, to those who wish to actually perform a Delphi study.

Forecasting methods can be considered as being of two types - quantitative and qualitative. The quantitative methods have been the mainstay of planners for many years. Extrapolations and other manipulations of historical data, in some cases using quite sophisticated mathematical and statistical techniques, have been used extensively to establish the parameters of the future.

These purely quantitative forecasts, while relatively simple to perform, create a planning dilemma because:-

- In a world where basic values are changing at an accelerating pace substantial discontinuities can be expected which will make the future significantly different from the past.
- Projections of past trends are most useful when forecasting on a macro basis for the whole economy as minor conflicting trends tend to balance out, but are unreliable when predicting developments in a relatively small segment of the economy.

- Many nations are reaching a level of affluence and a degree of technological sophistication which allows deliberate choices to be made about what direction to take in the future.

For these reasons forecasts which include the qualitative opinions of humans, in addition to the usual quantitative estimates, are needed. It may, for example, make economic sense and be technologically possible to take a certain course of action, but for political, social, or environmental reasons, a quite different choice may be made. A Delphi study is an attempt to bring the knowledge and intuition of a group of qualified individuals to bear upon the future possibilities in a given field. The group or panel opinion is sought concerning what likely will happen when all factors - social, technological, economic and political - are taken into account. The Delphi technique, therefore, is an organized way of arriving at a qualitative forecast which may vary considerably from past trends. As the world grows more affluent and more options become feasible, this method of developing consensus about the future will become an increasingly important weapon in the planners arsenal.

Delphi was developed at the RAND Corporation by Theodore Gordon, Olaf Helmer and Norman Dalkey. Essentially it is a method for developing and improving consensus on any numerical estimate and, while it is most often used in forecasting, the Delphi method can also be utilized to test and develop group agreement about contemporary questions. With the Delphi method a panel of individuals, who are knowledgeable in the area to be explored, forecast likely developments in that field. The distinguishing feature of the technique is that it relies upon several rounds of questionnaires to explore the views of the panel with statistical data and summaries of views of the other panelists being fed back to individual participants after each round. This procedure creates a modified form of group interaction and exchange of

views. It, at the same time, removes many of the counter-productive elements present in fact to face meetings such as the effect of status, group pressures to conform to majority opinion, the persuasive and dominant personality whose views may be quite wrong, and many other interpersonal variables. It replaces these influences with some distinctly different characteristics - notably anonymity, iteration and controlled feedback - all of which tend to foster calm, contemplative consideration of the issues. Reasonably comprehensive evaluations of the technique have found that it is a signifl-cant improvement over normal group meetings both in terms of arriving at a group consensus and in improving the accuracy of group forecasts. Delphi forecasting studies have been permformed by many corporations, institutions and government agencies concerned with social change, technological change in Education, Medicine, war prevention techniques and many other subjects.

As the prime purpose of this paper is to provide pragmatic guidance on how to design, write and administer a Delphi study, actual examples of questions and design layout will be used from three studies. These are

Frank J. Doyle and Daniel Z. Goodwill - An Exploration of the Future in Educational Technology, Bell Canada Report, January 1971.

Frank J. Doyle and Daniel Z. Goodwill - An Exploration of the Future in Medical Technology, Bell Canada Report, March 1971.

Frank J. Doyle - Future Electrical Transmission and Distribution Technologies, Office of Science and Technology Report, August 1973.

These studies will be referred to as the educational, medical and electrical studies in this paper.

HOW TO DESIGN AND WRITE A DELPHI STUDY

The design steps are outlined below and are followed by a more detailed description of each point.

1. Document the Study Objectives and time horizons.
2. Define five to ten broad cohesive areas of study.
3. Research each of the broad study areas.
4. Develop a draft questionnaire.
5. Test and finalize the draft questionnaire.
6. Design the analytical system to analyse the results of each round.
7. Select and recruit the Delphi panelists.
8. Develop a study schedule and mail Round I.

1. Document the Study Objectives and Time Horizons

A clear description of the purpose of the study and the time period to be studied should be written. It is usually preferable to define a prime time-frame and a secondary time-frame to allow panelists to completely describe their opinions about the timing of events. As an example the fundamental forecasting period of a study might be 1980 - 1990 with a decreasing interest in the period from 1990 to 2000.

2. Partition the topic into a number of broad cohesive areas of study

The topic of the study should be segmented into between five and ten broad cohesive areas and these arranged in a logical sequence. The intent is to organize the study so that a rational flow of thought occurs as the participant works his way through the questionnaire.

The following is an example of the partitioned sub-topics taken from the educational study.

- Changing values in society
- The Evolution of School Design and the Changing Role of the Teacher.
- Computer Assisted Instruction Systems.
- Information Retrieval Systems
- Audio-Visual Retrieval Systems
- Terminals
- Technology in The Home
- General Trends in Education

There should be a deliberate attempt to lead the thinking of the respondent from a consideration of broad change in the society to more narrow changes in the particular area being explored - in the above study for example from a consideration of changing values in the whole society to estimates of the adoption of educational technology in the home and school.

3. Research Each of the Broad Study Areas

Research should be performed on the overall study topic and on each of the sub-divided study areas to determine:

- What are the key parameters for the total study and for each study section?
- What has been happening and what might occur in the future to these parameters?
- Where might present trends lead in the time-frame of the study?
- What are the bizarre possibilities that could make an important difference in each section?

4. Develop a Draft Questionnaire

a) Two basic decisions must be made before the questionnaire can be written - these are:

- Should the study be designed to foster interaction among the participants and if so, to what extent?
- How many questionnaire rounds should be planned?

Interaction in a Delphi study is maximized when written comments and answers are given by the participants and minimized when structured, limited-choice responses, without accompanying comments, are sought. Table 1 shows the relationship between interactiveness, effectiveness and number of rounds and illustrates (in the opinion of the writer at least) that a three round, highly interactive, structured questionnaire with comments is the optimum design for most Delphi studies.

TABLE 1

Level of Interaction in study	Number of Rounds and degree of structuring in questions.	Level of effectiveness of study approach
Very High	Four of five Rounds - no structure in Round I. Participants establish questions in Round I.	Very Low
High	Three Rounds - Uses structured questions in addition to many comments and some open questions	High
Low	Totally structured with no comments.	Low

The fundamental reason for using the Delphi technique is to obtain qualitative human opinion about the occurrence of relevant events. This means that comments are critically important to the study - in fact the reason why a panelist believes an event will or will not occur may add more to the study than an estimate of timing. A completely unstructured study, where the panelists are requested to pose the questions in Round I, is an interesting exercise in futility which usually leads nowhere and requires four of five rounds to complete.

The three round, highly interactive study should be structured as follows:-

Round I

Poses relevant possible developments and requests the panelists to decide upon their likelihood and/or timing, desirability etc. Comments on why an event will or will not occur (or why it is or is not important) should be sought as well. Participants should be asked to suggest other possible developments and questions to be included in Round II.

Round II

A statistical summary of the opinion of the entire panel, usually showing the interquartile range and the median, should be included for each answer as well as a summary of the comments on each question. Panelists whose answers are outside the interquartile range should be asked to give reasons for their opinions. New questions suggested by participants would be included in Round II.

Round III

Round II opinions should be summarized as above and dissenting comments emphasized. Panelists should be ased to consider the statistical summaries and the reasons for the dissenting views and to arrive at a final answer to each question.

b) Write a study preamble, including a broad neutral scenario of likely change for the period of the study. This scenario should not be too explicit or detailed as this might constrain speculative thought. It could contain data such as the estimated GNP per Capita, population size, a description of major social trends etc., for the period to be examined.

Write a short preamble to each section of the study describing the main characteristics of the topic in that section and perhaps speculating about possible future changes. An example, taken from the electrical study is below.

"The Electrical Power Industry in most developed countries is faced with conflicting pressures. On one hand rapidly increasing demand requires that more and more power be delivered to industrial and residential customers while at the same time environmental pressures are growing to limit the bulk transmission structures designed to carry the required power. The dilemma presented by these conflicting demands constitutes a major technological challenge to the designers of future transmission systems. Potential solutions include:

- Underground Transmission of bulk power.
- The development of UHV transmission systems which carry more power on structures.
- The elimination of conventional towers by new approaches such as Micro-wave transmission.
- The development of on-site fuel cells.

All of these broad technological possibilities except the fuel cell, would require a number of developments in system components in the areas such as power generation, conduction, insulation, etc. The prime purpose of this section is to explore the views of the panel on the likely direction of the leading-edge technologies in these areas and their components over the next ten to twenty years."

c) Write the questions for each section

While each Delphi study is unique the variables being examined are often similar. The most common are:

- Timing of possible developments
- Desirability of possible developments
- Importance of possible developments
- Probability of event over time

Some examples are below from the educational, medical and electrical studies.

1) "Please show your estimate of the probability and importance of the following events.

Medical Personnel	Probability of Event				Importance of Event			
	Highly Likely	Quite Likely	As Likely as not	Unlikely	Very Important	Important	Little Importance	Un Important
General Internist taking the place of the G.P.								
Physicians practicing only outpatient or inpatient medicine								

2) When do you expect conventional power delivery systems will be replaced by fuel cells in 10% of industrial locations? Please show the probability of occurrence for each time period.

	Expected Period of Adoption						
Fuel Cells in Industrial Locations	72-75	76-80	81-85	86-90	91-2000	Later	Level of Expertise
	0	0	5	10	30	55	3

The above answer indicates that this panelist expects this event to occur, that he estimates a 55% chance it will not occur before 2000, a 5% chance it will occur by 1985, a 10% probability between 1986-1990 etc. This answer also illustrates another analytical aid - the self-rating of expertise. This study asked respondents to estimate their level of expertise for each question using the following coding.

1 - Intelligent guess

2 - Knowledgeable

3 - Expert

This allows the results to be analysed and reported in a variety of ways. For example, in Rounds II and III median scores for the entire panel opinion and for the opinion of experts could be shown as additional information for participants.

3) A second approach to timing (and one that is much simpler to analyse) requests the panelist to estimate the most likely time period that an event will occur in as below from the Educational study.

When, if ever, do you expect that a significant percentage (20%) of learning and study will be carried out in the home using sophisticated technology?

School Level	70-75	76-80	81-85	86-90	91-99	Later	Never
Primary							
Secondary							
Post Secondary							

4) An Example of a Broad General Question

In both the Medical and Educational studies a general question on changing values in society was used as the first question in the study to start participants thinking about broad, general change and its likely future direction. The question used in these studies (preceded by a suitable preamble) is below:

Value Changes in North American Society					
Value	Significant Increase	Slight Increase	No Change	Slight Decrease	Significant Increase
Traditionalism					
Hard work as a Virtue					
Rewarding work as a Virtue					
Self Expression					
Acceptance of Change					
Involvement in Society					
Participation in Decision making					
Authoritarianism					
Individualism					
Materialism					

5) The Time/Pattern Table

Some situations allow the design of a Table showing a number of possible related events that may occur sequentially over a period of time. An example from the Educational study is below:

The Evolution of School Design								
Trend		70-75	76-80	81-85	86-90	91-99	Later	Never
1.	Traditional Classroom	—M—						
2.	Some Classrooms - Some Communal shared areas	M——						
3.	No Classrooms Communal shared areas			——	M ——	——	——	
4.	Small open area community schools supplemented by community resources			——	M ——	——		
5.	Schools used for occasional seminars. Most learning outside schools					M ——	——	

The "M" indicates the median answer and the horizontal line the interquartile range of the panel's answer to this question. The fact that all of these events are in one table assists the respondents to think in logical patterns and to relate mentally, events that affect one another.

6) The Inclusive Table

This Table shows a number of possible future events that taken together include all of the possibilities and therefore add up to 100% as in the question below from the electrical study.

"AC or DC Transmission of Bulk Power

Some authorities believe that DC transmission will be used more in the future This may be particularly true for transmissions over long distances or for underground systems Please show in the following table the percentage of bulk power transmission you expect will use AC and DC in the future for underground transmission".

Underground Systems

	Level of Expertise 1, 2 or 3	Expected Period of Adoption				
		72-75	76-80	81-85	86-90	The Year 2000
DC						
AC						
Total		100%	100%	100%	100%	100%

7) In each section at least one open question should be used to encourage panelists to sum up their feelings on the section in written form An example - "What do you consider to be the most important technological development required for Ultra High Voltage transmission by 1985?" The open question should usually be the last one in the section.

8) In most studies a number of relevant events will not fit readily into one of the study sections. The best way to handle this problem is to create a miscellaneous section - usually consisting of a table listing a number of unrelated possible developments. Some examples from the miscellaneous section of the medical study are shown on the following page.

Some Technological Developments

	70 - 75	76 - 80	81 - 85	86-90	91 - 99	Later	Never
Physicians will perform many functions automatically from their offices such as: 1. Obtain room reservations in hospitals immediately.		M—					
2. Change hospitalized patients diets		——	M	——			
3. Evaluate hospitalized patients progress by remote audio-visual contact and adjust treatment if necessary			——	M			
4. Make home "visits" by using audio-visual services			——	M	——		
Physiological monitoring during operations will provide earlier crisis warning	M—	——					
Dependable artificial hearts will be developed which will last 10 years			—	M	—		
Several hospitals will share a large computer for administration, diet analysis, billing, statistical control on patient stays and bed utilization as well as a variety of other purposes		——	M				

The examples above are some of the common question formats used in Delphi. Many others are of course, possible. In general the objective should be clarity and simplicity. Another important point is that questions should not ask for answers which are too fine-grained - it is not realistic to expect panelists to give extremely detailed opinions about the long term future - even though the researcher is always tempted to obtain responses that are as explicit as possible.

5. Test and Finalize the Draft Questionnaire

A Delphi study is a difficult exercise in written communication. No matter how much effort is devoted to achieving clarity and understanding in the questionnaire, confusion about the precise meaning of some questions will exist with some of the participants. Between six and ten individuals, who might be typical of panel members, should be asked to complete the draft questionnaire and to list any ambiguities. These should be eliminated and the questionnaire finalized.

6. Design an Analytical System to analyse the study results.

The complexity of the anlaytical system can range from an elaborate computer program to a simple arrangement of categorizing and cataloguing comments and choices, depending upon the study design, the type of question and the number of panelists. If it is decided to request the panelist to estimate the probability of occurrence of events over several periods of time, a relatively elaborate system is required. An example is shown in Appendix I which is an excerpt from the Electrical study instructions in Round II. If respondents are asked to select the most likely period of adaption it is much simpler to analyse the results. All comments should be summarized and recorded for each question.

7. Selection of Panelists

Delphi is a technique which links together the knowledge and intuition of a group of individuals to develop a consensus on the issues being explored (or define areas where a consensus is not possible). The quality of the results of a study, therefore, depends entirely upon the calibre of the participants. Panelists should be recruited who have a comprehensive knowledge of the areas under study and who are likely to be broad, intuitive thinkers. The panel as a whole should represent a cross-section of viewpoints and backgrounds so that all relevant perspectives in the field being studied are included. For example, in the Education study individuals with backgrounds in the following areas were on the panel.

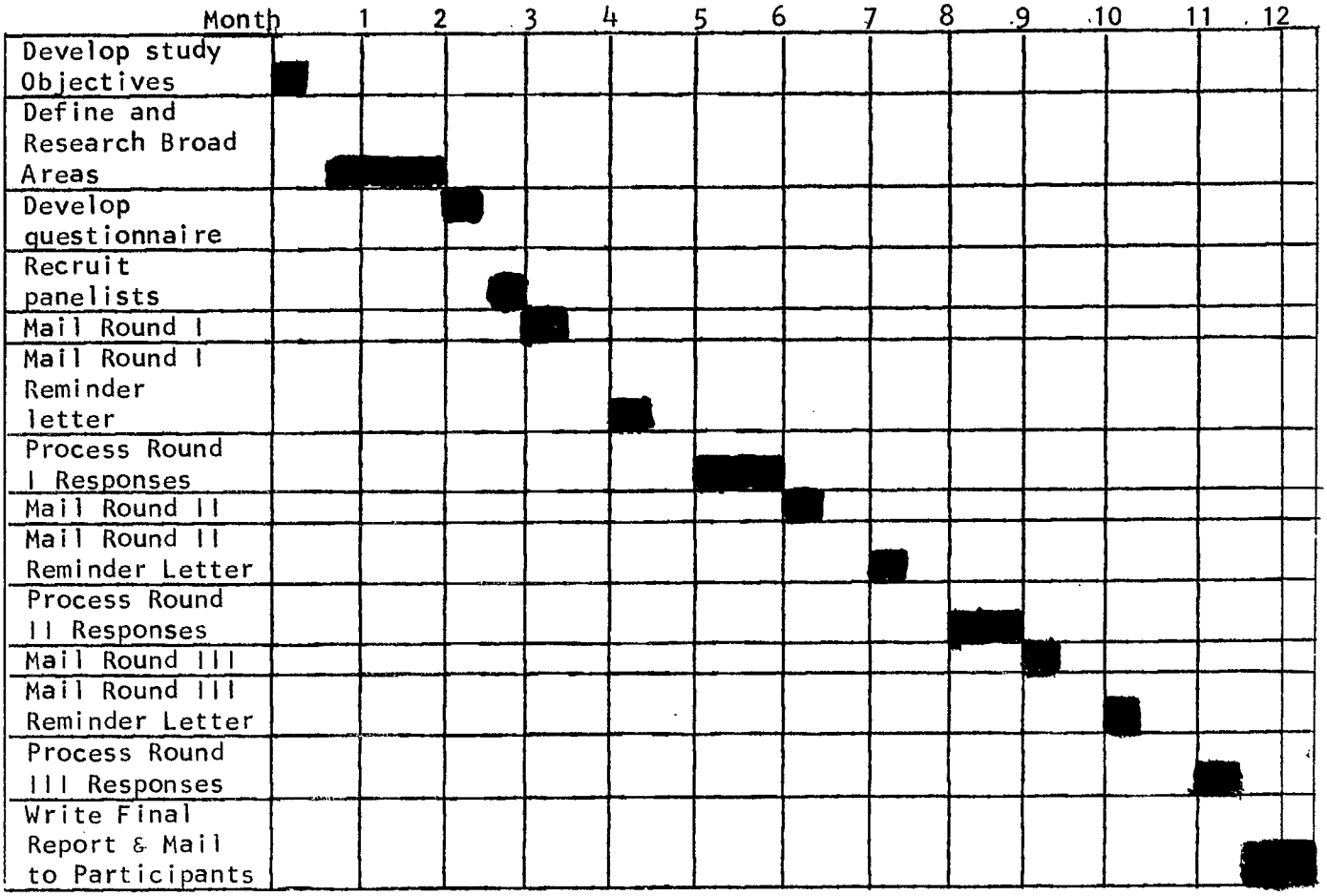
- Audio-Visual experts
- Librarians
- Researchers in Educational Methods
- Researchers in Computer Applications for Education
- School Administrators
- Principals, Deans, Professors, Teachers
- Educational Administrators in Government
- Educational Consultants
- Teachers Associations

Once an analysis has been made of the various "schools of thought", knowledgeable persons in the field should be consulted to obtain the names of the best prospective panelists to represent these various viewpoints in the study. If it can be arranged, panelists should be asked to participate in the study by someone they know, not by the researcher. Experience has shown that the likelihood of respondents refusing to participate or dropping out in the course of the study is minimized by this approach.

Many studies are conducted with panels of between ten and twenty respondents. While this small panel does simplufy the administration and analysis of the study, in most cases it probably results in too narrow a range of opinion as it can not adequately represent all viewpoints. Studies in most fields require between forty and sixty panelists.

8. Develop a Study Schedule

A typical schedule is shown below:



A typical three-round study will require approximately one year to complete and will utilize about eight man months in time. The "reminder letter" shown in the schedule is a polite request to complete and return the outstanding questionnaire. Processing can begin when seventy or eighty per cent of questionnaires have been received. The dropout rate in Delphi studies varies from 10% to 80%. The researcher must always be aware of the fact that panelists may become bored or lose interest and make every effort to foster comment and involvement. In the three studies used as examples here the dropout rate was held to about 20%.

Summary

The Delphi technique is a flexible tool which can be used for forecasting, consensus building on contemporary issues, as an aid to policy formulation and in many other situations where collective human opinion is required. There are no hard and fast rules in the application of the technique and experimentation designed to cope with particular problems is a necessity in each study. In fact, the number of possible variations is limited only by the ingenuity of the designer. The purpose of this paper is not to set limits or rules but rather to describe in some detail the basic design considerations and to give neophyte practitioners a running start in their experimentation and development of the technique.

Excerpt from the Round II Instructions in the Electrical Study

The following summary statistics have been calculated as appropriate:

(i) Mean expected date and probability of occurrence

For a question requesting probability (percentage) responses, the answers were all converted to distributions totalling 100%. If the sum of entries across the row totalled less than 100%, the remainder was called a "later/never" prediction. Then an average was calculated for each time period and the later/never column.

First, the average of the later/never column, after normalization, represents the "probability of nonoccurrence" during the indicated time periods. Taking the difference between 100% and this number gives the "probability of occurrence" (represented in this discussion as p (occurrence)).

Next, a mean expected date of occurrence (represented as mED) was calculated under the assumption that the event would occur some time during the indicated time frame. In other words, the sum of the average responses for each of the time periods was normalized to 100%, and an expected date calculated by multiplying the average normalized probability for each period by the period's mid-date and summing the results.

The significance of these mean expected dates and probabilities of occurrence is that they represent the overall average response. The mean expected date does not indicate the spread of answers but this will be indicated as described below. Also, averages can be misleading since they can be easily biased.

The probability of occurrence number is very important, however. It is a measure of confidence for the mean expected date. A low probability of occurrence is indicative of the fact that many of the respondents do not feel that the event will occur - and the mean expected date and other statistics to follow should be taken in this light.

A mED and p (occurrence) was calculated for the overall group and for those rating themselves a 3 in expertise alone.

(ii) Median expected date, interquartile range, and range

An expected date was also calculated for each response given that the event would occur (if probabilities were requested). The median of these dates has been called the median expected date (represented as MED). Two medians were determined, one in the normal fashion and one weighting the responses by the indicated level of expertise. In the analysis, the unweighted median is shown as a triangle (▲), underneath the approximate date on a scale of years. The weighted median is shown by a star (*).

The interquartile range is the middle quarter of the expected dates by respondent and is shown by the bar on the scale. The range indicates the extreme answers.

The significance of the MED and interquartile range is that the answer minimizes the effect of respondents at either extreme (very optimistic or pessimistic). Remember, however, that these predictions assume that the event will occur during the time frame under study. The probability of occurrence value is also a measure of the confidence in these predictions.

(iii) Average

Where the responses did not ask for probabilities, simple averages were calculated in the normal fashion.

In summary, the statistics calculated are:

Statistic	Significance
1. Mean expected date (mED) overall and experts only	Overall average of responses converted to an expected date given that the event will occur.
2. Probability of occurrence	Confidence in the MED and mED, i.e. the probability that the event will, in fact, occur.
3. Median expected date (MED) unweighted and weighted	The middle responses, unbiased by the extremes, given that the event will occur.

Please consider these statistics and associated comments when formulating your new response. Although tedious, this is the essence of the DELPHI technique.

Thank you.