THE DELPHI METHOD

by

Theodore J. Gordon

Acknowledgments

Some contents of this report have been taken, in some cases verbatim, from internal papers of The Futures Group with their permission. These papers were written by John G. Stover, Theodore J. Gordon, and others describing the Delphi method and its applications. The managing editor also gratefully acknowledges the contributions of reviewers of the draft of this paper: Dr. Ian Miles of The Programme of Policy Research in Engineering Science and Technology, in the United Kingdom; Dr. Brian Free of the Futures Environment Council of Alberta, Canada; Dr. Mika Mannermaa of the Futures Research Centre at Turku School of Economics, Turku, Finland; Dr. Harold A. Linstone of Portland State University, United States; and Dr. Peter Bishop of the University of Houston, in the United States. And finally, special thanks to Elizabeth Florescu and Neda Zawahri for project support, Barry Bluestein for research and computer operations and Sheila Harty for editing.

I. HISTORY OF THE METHOD

The modern renaissance of futures research began with the Delphi technique at RAND, the Santa Monica, California, "think tank" in the early 1960s. The questions of RAND thinkers, at the time, primarily dealt with the military potential of future technology and potential political issues and their resolution. The forecasting approaches that could be used in such applications were quite limited and included simulation gaming (individuals acting out the parts of nations or political factions) and genius forecasting (a single expert or expert panel addressing the issues of concern). Quantitative simulation modeling was quite primitive, and computers that would ultimately make such quantitative techniques practical, were not yet capable enough.

The RAND researchers explored the use of expert panels to address forecasting issues. Their reasoning went something like this: experts, particularly when they agree, are more likely than non-experts to be correct about questions in their field. However, they found that bringing experts together in a conference room introduces factors that may have little to do with the issue at hand. For example, the loudest voice rather than the soundest argument may carry the day; or, a person may be reluctant to abandon a previously stated opinion in front of his peers. As with normal thinkers, the give-and-take of such face-to-face confrontations often gets in the way of a true debate.

One of the little known in-house research projects undertaken by RAND at the time involved combining opinions of horse-racing handicappers. These people, after all, are supposedly experts in their field. Furthermore, their opinions about the future (the outcome of horse races) are published daily and can be checked against reality within 24 hours. So a project was implemented to determine just how to combine horse-race forecasts by different experts to improve the likelihood that the composite opinion was better than any single expert.

The work on the Delphi method followed. Olaf Helmer, Nicholas Rescher, Norman Dalkey, and others at RAND developed the Delphi method, which was designed to remove conference room impediments to a true expert consensus. The name, of course, was drawn (humorously, they thought) from the site of the Greek oracle at Delphi where necromancers foretold the future using hallucinogenic vapors and animal entrails. They began from a philosophical base and asked initially, "just how much could be known about the future?" (Helmer and Rescher, 1959)

The Delphi method was designed to encourage a true debate, independent of personalities. Anonymity was required in the sense that no one knew who else was participating. Further, to eliminate the force of oratory and pedagogy, the reasons given for extreme opinions were synthesized by the researchers to give them all equal "weight" and then fed back to the group as a whole for further analysis. These aspects: anonymity and feedback, represent the two irreducible elements of the Delphi method.

This general approach has been used thousands of times since the first published Delphi study, *Report on a Long-Range Forecast* by Gordon and Helmer. This RAND "best seller," published in 1964, contained forecasts of scientific and technological breakthroughs through 2000 and beyond; the 82 panelists who contributed included Isaac Asimov, Arthur Clarke, Bertrand de

Jouvenel, Ithiel de Sola Pool, Dennis Gabor, Peter Goldmark, Harold Guetzkow, and William Pickering, to name a few. (Gordon and Helmer, 1964)

The temptation to review all of the forecasts with the advantage of hindsight is great (one such review was done by Amant in 1970), but a cursory review shows many forecasts that were on target, such as:

- economically useful desalination of sea water
- oral contraceptives
- advent of ultra light materials
- automated language translation
- transplanting organs
- more reliable weather forecasts
- centralized data banks
- artificial organs
- x-ray lasers
- psychotropic drugs
- self-replicating molecules
- synthetic proteins
- feasibility of control over hereditary defects

There were big misses, too, including:

- controlled thermonuclear power
- biochemical general immunization
- limited weather control
- world population by 2000 less than 6 billion
- manned landing on Mars

Nevertheless, the study and, more generally, RAND's interest in developing systematic methods for forecasting apparently legitimized the field. Since this first Delphi, literally thousands have been performed on topics as wide ranging as the future of religion and the family, to space exploration.

The method continues to be used extensively. In September of 2008, a review was made of the Scopus data base (which includes articles from 15,000 peer-reviewed professional journals from 4,000 publishers, proceedings papers, and trade publications) and 105 publications were identified in response to the search terms “Delphi study.” By far, the largest field of study which employed the method was health sciences. An incomplete Google search on “Delphi Studies 2008” produced research reports and study descriptions that included the following topics, all published in the first nine months of 2008, listed here to indicate the extremely broad scope of recent applications:

- Asthma indicators (refining the list)
- Back pain definitions
- Chest x-rays
- Chronic anorexia nervosa
- Clinical information systems
- Electronic records management
- Gerontology
- GIS science
- Health education
- Health promotion
- HIPAA compliance on the Battlefield Medical Evacuation process.
- Library science
- Lowering CO emissions from diesels
- Mathematical problem solving (in primary school)
- Noxious stimuli (neonatal)
- Nursing education
- Palliative medicine
- Parenting and child health
- Physiotherapy
- Poverty alleviation
- Power sector- Kerala State, India
- Psychiatry
- Public relations research priorities
- Severe Acute Respiratory Syndrome
- Technology management

Two studies from this set are described here as recent examples of the use of the method:

> Pare, G, C.Sicotte, M. Jaana, and D. Girouard, "Prioritizing clinical information system project risk factors: A Delphi study," reporting on risks associated with implementing a clinical information system. The research involved a Delphi among 21 experts involved in Canadian CIS projects to identify risk factors of clinical information systems.
>
> Dionne, C.E., et al., "A consensus approach toward the standardization of back pain definitions for use in prevalence studies, " in which a Delphi study was performed with 28 experts in back pain research from 12 countries. This work was designed to:
>
> > "identify standardized definitions of low back pain that could be consistently used by investigators in prevalence studies to provide comparable data….[and] the study produced definitions [that could] provide standards that may improve future comparisons of low back pain prevalence figures by person, place and time characteristics, and offer opportunities for statistical summaries."

II. DESCRIPTION OF THE METHOD

This description will be illustrated by a hypothetical example: suppose we want to establish a forecast for the date by which a manned Mars landing would occur. With the Delphi method, experts from the required disciplines are first identified and asked to participate in the inquiry. For this example, the experts might include rocket scientists, geologists, and bioscientists who are experts on the planet, planners from space agencies that might be involved in such a project, and others who expressed an opinion that manned landings might be a bad idea. During the initial contact, the nominated persons are told about the Delphi and invited to participate. They are assured of anonymity in the sense that none of their statements will be attributed to them by name.

The questions are refined by the researchers and pursued through a number of sequential questionnaires. In the first questionnaire, participants might be asked to provide their judgment on a date at which a manned landing might take place. The analysis would identify the range of opinions about the date. In a second questionnaire, the range would be presented to the group, and persons holding opinions at the extremes of the range would be asked to reassess their opinion in view of the group's range and provide reasons for their positions. For example, a reason for a late manned landing might be that robot landers can do all that men can do, so little reason exists to spend the money for human on-site involvement.

These reasons would be synthesized by the researchers at the end of round two; the synthesized reasons would form the basis for the third questionnaire. In this third questionnaire, the new group judgment on a date would be presented to the participants, along with reasons for the extreme opinions. Each member of the group would be asked to reassess his or her position in view of the reasons presented. They might also be asked to refute, if appropriate, the extreme reasons with any facts at their disposal. For example, someone might attempt to refute the robot argument by saying that human presence will be required to perform a certain class of engineering activity or to construct habitats for later colonization.

In a fourth and final round, these arguments would be presented, along with the evolving group consensus, and a reassessment requested.

In a sense, the Delphi method is a controlled debate. The reasons for extreme opinions are made explicit, fed back coolly and without anger or rancor. More often than not, groups of experts move toward consensus; but even when this does not occur, the reasons for disparate positions become crystal clear. Planners reviewing this material can make judgments based on these reasons and their own knowledge and goals.

Because the number of respondents is usually small, Delphis do not (and are not intended to) produce statistically significant results; in other words, the results provided by any panel do not predict the response of a larger population or even a different Delphi panel. They represent the synthesis of opinion of the particular group, no more, no less.

The value of the Delphi method rests with the ideas it generates, both those that evoke consensus

and those that do not. The arguments for the extreme positions also represent a useful product.

Questions included in a Delphi may be of any sorts that involve judgment, including, for example, the size of a future market, whether or not the CEO should receive a raise, or the proper policy to achieve a goal. In planning applications, the questions generally are of three types.

> *forecasts on the occurrence of future developments.* Forecasts of future developments call for answers about when an event is expected to occur or about the future value of some parameter;
>
> *desirability of some future state.* Questions dealing with desirability ask for judgments about whether an event **ought** to occur, and the basis for the recommendation; and
>
> *the means for achieving or avoiding a future state.* Questions dealing with policy involve the traditional reporter's questions about implementation: who, what, when, where, and how much? But to this set we must add: to what end. In other words, questions about policy ought to be linked closely to the objectives sought and the likelihood that any policy will, in fact, accomplish its intended goals.

These three types of questions may require different kinds of experts. The likelihood questions may involve hands-on experience and intimate knowledge of the frontiers of research. The desirability questions may involve a moral, political, or social dimension quite distinct from the disciplinary expertise involved in judging likelihood. The policy question may involve knowledge of the art of the possible.

In some modern applications of Delphi:

1. The questions relate to the value of independent variables that are used in **quantitative simulation models**. In this application, a consensus is not required; rather, if disagreement exists about the value of any variable, the extremes can be tested in quantitative models to determine whether or not the difference has any important significance.

2. **In-depth interviews** with experts have been used with great success as an alternative to questionnaires. In this approach, the same kinds of experts are first identified, invited to participate, assured of their anonymity, and, in most instances, promised a report based on the interview sequence. Appointments are made at the convenience of the interviewees. Interview protocols are prepared and tested to elicit judgments. High-level staff members, familiar with the study's objectives, act as interviewers. Feedback can be introduced if two rounds of interviews are employed; however, single-round studies are used more frequently. In these, "feed-forward" is often employed, presenting to respondents information about emerging consensus derived from the prior interviews. True, this process introduces differences among the various interviews, but remember that the exercise is not designed to be statistically significant but rather to elicit ideas that can be important to subsequent analyses. Expert in-depth interviews are an excellent means of obtaining such ideas.

An advantage of one-on-one interviews is that they provide flexibility, which is absent in

questionnaires. For example, an interview provides the opportunity to probe the reasons behind the forecasts, to search for biases in the forecasts, and to follow up on unexpected hints dropped by the interviewees.

3. For some applications, **group meetings** among experts have now become practical. Delphi had its birth in concern about spurious factors that intrude in face-to-face meetings among experts. New technology can minimize these factors. Some Delphi-like studies have been performed on-line (Shota, 1993) and experiments date back to the 1970's when Murray Turoff experimented with early computer-based communications to link experts together in networks. (Turoff, 1972). Most recently he and his colleagues have described a Social Decision Support System in which large groups of people (thousands) interact and vote dynamically (can change votes as in Delphi) on social issues. (Turoff, Hiltz, Cho, Li, and Wang, 2002)

Conference room voting machines have proved useful. The Consensor (Applied Future, Westport, CT) built and rents one such machine. Others include OptionFinder (Minneapolis) and the PC Voter (The Futures Group). Typically, these machines provide each participant with a small terminal, which is connected through a serial circuit to a personal computer. In the PC Voter, each small terminal has two knobs. The first knob allows the user to provide quantitative judgments about a question posed by the meeting's moderator for example: "What is the probability that limited weather control will be available by 2020?" Using the second knob, participants can provide a percentage assessment of confidence in their answer. The computer's software integrates the answers of the experts at the meeting, discounting those who have low confidence in their answers, and provides on a display screen a histogram showing the distribution of the group's opinions. Anonymity is preserved because the inputs are private and unseen by others; the display provides feedback. It is true that anonymity is lost in any discussion of the results, but discussion is an option of the participants. This approach has been found quite useful when quick results are needed and the issue is, "hot"; that is, the topic of the study is likely to evoke strong emotional responses. An example of a "hot" issue is the future of executive compensation among the executives and compensation committee members of a corporation.

4. Several recent studies have placed their questionnaires on-line to reduce the communications delays and burdens. These studies have substituted the communications tools available via Internet for the old paper forms that were transmitted by fax, postal mail, or more recently e-mail, Nevertheless, the studies still employed sequential questionnaires and preserved the Delphi principles of anonymity and feedback,

5. Recently, the Millennium Project of the World Federation of United Nations Associations has developed "Real Time Delphi" a method described in detail elsewhere on this CD ROM. Briefly, in this approach a Delphi-like questionnaire is accessed by respondents from an on-line Internet web site. The answers are typed in and on submittal are fed to a database on the server. The average response is computed and fed back to the participants who are also asked for the reasons behind their answers. Each respondent can see the reasons given by all respondents and may at any time change their answers. There is no explicit second round. When the respondent comes back to the study in a minute or a day, the original input form with their prior answers is presented to him or her. Of course, by then others may have contributed judgments, the averages

or medians may have changed and given this new information, the respondent may wish to revise their prior answers. In this way the Delphi requirements of anonymity and feedback are met and the process, once underway, yields the distribution of the group's responses and reasons for the extreme positions. The process can be synchronous or asynchronous, and if implemented on an Internet site can easily involve worldwide panels..

III. HOW TO DO IT

The key to a successful Delphi study lies in the selection of participants. Since the results of a Delphi depend on the knowledge and cooperation of the panelists, it is essential to include persons who are likely to contribute valuable ideas. In a statistically based study, such as a public opinion poll, participants are assumed to be representative of a larger population; in Delphi, non-representative, knowledgeable persons are needed. So the first problem is how to select potential participants.

Knowledgeable persons are usually identified through literature searches for who has published on the subject under study, recommendations from institutions (e.g., the World Future Society) and other experts in a process known as "daisy chaining." Literature searches necessarily result in lists of people who have published on the topic of interest; this approach misses people who may have something to contribute but have not published. Recommendations from institutions suffer from the same shortcoming: the recommendations are limited to only those who are known to the institutions. "Daisy chaining" has the potential disadvantage of identifying cliques. One possibility that helps ensure that the required skills are represented is to form a matrix in which the required skills are listed. In the case of the Mars landing question, for example, the columns might be: solar system planetary geology, rocket system design, robotics, data collection, telemetry, etc. The rows of the matrix consist of the names of the prospective participants. The cells are checked to indicate the "coverage" that the nominees are expected to provide.

As for "unknown" people who are outside of the normal lines of communication but who may be able to contribute new and innovative ideas, here are some suggestions:

> use bulletin boards so that contributors who have something to say in this informal environment may identify themselves;
>
> get recommendations from university professors about bright students;
>
> advertise for participants and qualified applicants through preliminary Delphi's.

Most studies use panels of 15 to 35 people, although in some applications hundreds, even thousands of respondents have been involved.. The length of the invitation list should anticipate an acceptance rate of between 35 and 75 percent.

Once the list of nominees is formed, each person should be contacted individually. Form letters should not be used. The initial contact may be by telephone, but letters should confirm the

invitation. These letters should contain a description of the project, its objectives, the number of rounds to be included (or the time commitment anticipated), the promise of anonymity, and, if appropriate, a confirmation of the panelist's acceptance.

The next step is to formulate the questions. The questions must be sharp and answerable. A small panel might be used to help formulate the questions. For example, say our question is a policy Delphi on nuclear proliferation and centers on the question:

> What effective and practical means exist to prevent proliferation of nuclear weapons?

The question would be sent to a small expert panel (10 persons or so) working with the research team. Essay-type answers would be permitted at this stage. The answers might include:

> Military intervention, greatly enhanced customs operations, new nonproliferation treaties, space surveillance, etc.

The research staff would collate the answers and form a non-essay questionnaire. The question might be:

> The following means have been suggested for preventing nuclear proliferation. Please rate each of these on a scale of 1 to 10, with 1 being the least effective and practical and 10 being the most effective and practical. Add other suggestions to the list if you have an idea that might be as effective and practical as those listed. If you think any of these means are extremely good or, at the other extreme, counterproductive, please provide your reasons.

This non-essay question would serve as the basis for the first questionnaire. The questionnaire first would be tested, perhaps using the small advisory panel. The testing would include actually filling in the questionnaire. This test is designed to find flaws in the way the questions are asked and to find any lurking possibilities for misinterpretation.

Once qualified in this way, the questionnaire would be sent to the participants. The cover letter would remind the participants about the objectives of the study, establish the schedule for the response, and include the return address. The media that are practical for transmitting the questionnaires and responses are postal mail, fax, and e-mail.

Our experience indicates that a response rate from 40 to 75 percent of the participants can be anticipated. Except for Real Time Delphi, the turnaround time is on the order of weeks, no matter what the medium of communications is.

In this example, respondents would provide two numbers for each option, the first depicting their judgments about effectiveness and the second, practicality of the actions suggested. They would be asked to provide reasons for their judgments and to suggest means for improving effectiveness and practicality as well as to state other alternatives they might have in mind. .

The research staff would collate the results. A feedback round would be used to present the

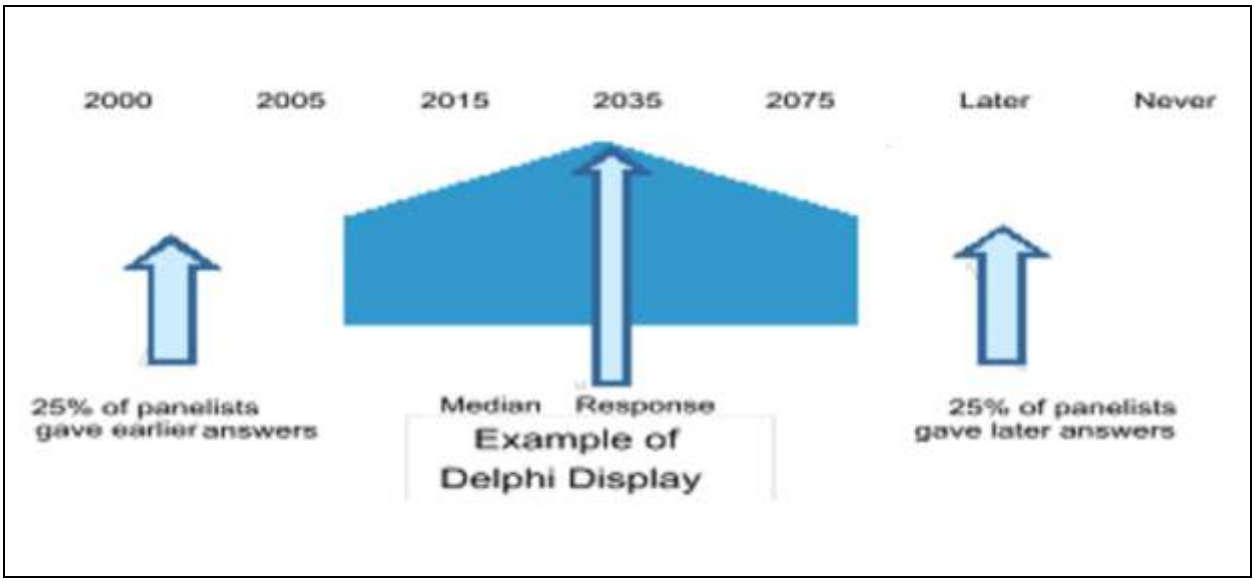
results, the reasons for the extreme positions, and a call for reassessment. The reasons, in Delphi style, would be placed in front of the participants who answered the first round. The respondents would be asked to reconsider their former answers in view of the reasons for the extreme opinions.

Questions that have quantitative answers or "check the box" responses are always easier to use. When the questions are in this form, they can be collated using spreadsheet software. In fact, in one application, the questionnaire itself was on a spreadsheet transmitted by e-mail, and responses were entered directly on the spreadsheet which was returned to the research team as a e-mail attachment. However, it is not always possible to use this approach or desirable to ask questions in this form. For example, we may want suggestions from the panel about policies for diminishing population growth rate: here, the essence is in the detail that the panel might provide. In addition, even when quantitative rather than narrative answers are called for, notes from the panelists may contain the most valuable information: references to other people, past experiences, uncertainty, pointing to data, etc. For this reason, all questionnaires must be read and digested by senior members of the study team.

Phrasing of questions is important. A common mistake is to include two events in the same question. Example: "When will bicycles be made mostly of plastics and used in urban transportation?" Differences in the way people answer this question may stem not only from their perceptions about future uses of plastic in bicycles, but also from differences in their beliefs about when bicycles will likely be used for urban transportation. Even subtle changes in wording may affect answers. For this reason, it is desirable to conduct a test of the instrument, not only by obtaining answers from a test group, but also by discussing with that group their interpretations of the questions.

Kahneman and Tversky point out that the framing of a question often contains seeds of the answers it will provoke. Presented with the choice between two therapies for a fatal disease, simply stating the cure rate in terms of survival rather than mortality biases answers in that direction. To avoid framing bias one might test the same question asked in different ways and compare results.

The data from a Delphi can be displayed in several ways. The group judgment should be based on the median rather than the mean, since single extreme answers can "pull" the mean unrealistically. Furthermore, it is incumbent on the analyst to show the spread of opinion, which can be done graphically or numerically by showing a range of responses, often the interquartile range (the range that contains the answers of the middle 50 percent of the respondents). An example is shown below:



Special care in presenting results must be taken if the distribution is unusual, e.g. bimodal.

One quantitative feedback option offered in Real Time Delphi is a display of the percentage or number of responses sorted into quintile groups stretching from the minimum to the maximum response, as follows:

Avg	1st Group of Answers	2nd Group of Answers	3rd Group of Answers	4th Group of Answers	5th Group of Answers	Total Responses	Standard Deviation
	14	18	13	13	2		
46.58	10.0 to 28.0	28.0 to 46.0	46.0 to 64.0	64.0 to 82.0	82.0 to 100.0	60	21.5

IV STRENGTHS AND WEAKNESSES OF THE METHOD

Delphi studies are difficult to perform well. A great deal of attention must be given to the choice of participants; the questionnaires must be meticulously prepared and tested to avoid ambiguity. Multi-round studies require a great deal of time; inevitably, some participants will drop out during the process. Therefore, other techniques for collecting expert judgment are welcome. Many of these techniques are described elsewhere on this CD ROM.

The primary strength of Delphi is its ability to explore, coolly and objectively, issues that require judgment; a weakness of Delphi is the ease with which questions can be asked for which better techniques exist. At one extreme are questions about the future for which factual answers exist and thus require minimal judgment. For example: "Given continued trends in immigration, birth, and death rates, how many people over the age of 70 will be in the United States in 2050?" This question has a numerical answer that can and should be computed. Factual questions should

not be candidates for Delphi. However, the question of whether or not current population trends will continue is a matter of judgment and, therefore, appropriate for Delphi.

In short, Delphi is a powerful technique when used to seek answers to appropriate questions. Suppose, for example, that we want to forecast the future size of the market for a given product. A Delphi study might involve sales and marketing personnel, retailers, and experts in consumer preferences. These people might be asked for a direct estimate; to do so would require that they somehow integrate all factors affecting the market, such as pricing, changing fashion, competition, consumer spending, etc. Participants might focus on different factors and account for them differently; without a direct question about the market determinants, the size-determining factors would remain hidden. Therefore, a better technique would be to ask for the respondents to identify the factors important to future market size and to estimate probabilities and consequent effect on sales, were they to occur. Then, using this information and a quantitative technique, such as Trend Impact Analysis, the size of the future market could be estimated, not only in view of the past market trend but also in consideration of trend-changing factors uncovered by Delphi participants.

Finally, a weakness of the Delphi method is the time that it takes. A single round can easily require three weeks; a three-round Delphi is at least a three- to four-month affair, including preparation and analysis time.

Delphi has its critics. Fred Woudenberg, for example, reports that "the main claim of Delphi — to remove the negative effects of unstructured, direct interaction - cannot be substantiated." He found that Delphi does not produce more accurate answers than other methods, and that consensus occurs as a result of pressure brought on participants who have extreme opinions. These questions are well worth exploring in more depth. (Woudenberg, 1991; Kasten et al., 1993; and Rowe G. et al., 1991). Gene Rowe and George Wright (Rowe and Wright, 2001) reviewed the literature comparing Delphi to other structured group procedures such as groups that have been instructed to argue both the positive and negative side of a question, and groups that used a structured form of information exchange. They found that “Given the equivocal nature of these studies, … there appears to be no clear cut rationale for adopting any of these techniques in preference to Delphi.”

True, the Delphi method makes participants with extreme opinions work harder than others. If opinions are not strongly held, participants may switch positions rather than write reasons for their estimates. On the other hand, those with strong opinions state why. The original impetus behind the method was to seek consensus, because expert consensus was believed more likely to be accurate than an individual forecast. Dalkey (1970) tested this proposition using obscure questions with known answers and found weak support for the concept. Today, consensus is less important for many investigators than previously; now a useful product of the Delphi method is crystallization of reasons for disagreement. Furthermore, Delphi is now seen as no more or less than a systematic means of synthesizing the judgments of experts—the aggregate judgment representing a kind of composite expert composed, in the domain of interest, of the expertise of all participants.

Does the method produce an accurate view of the future? It is no more accurate, probably, than

any expert, single or composite. But suppose we wanted to form a scenario based on expert views of what might be possible. Or suppose we needed a judgment about whether or not we could mount a manned Mars mission and if so, how. Or suppose we wanted to explore the range of future events that could affect population growth or weaponry or war. No better way exists to collect and synthesize opinions than Delphi.

V. FRONTIERS OF THE METHOD

The central question on the Delphi frontier is "*How can questions be addressed to the persons most likely to answer them well?"* Here are some thoughts that were developed in the Phase 1 Millennium Feasibility Study (Gordon and Glenn, 1993).

Self-Rated Expertise. If the number of questions and participants is small, then one might say "let's have everybody answer all questions." This approach is essentially taken in most Delphis. If the answers of all participants are accorded equal weight, the analysis is easy: for questions that can be answered in numerical form, the answers are often simply displayed as an average (or in statistical terms, such as median and standard deviation). Since sample sizes are small, this kind of analysis can give the appearance of numerical respectability when none really exists. For nonnumeric questions, the participants' answers can be displayed as a spread of opinions.

Some researchers (e.g., Dalkey and Helmer) attempted to introduce a refinement: self- rated expertise. The reasoning went something like this: an expert ought to be believed more than a non-expert; therefore, we ought to give an expert's opinion more weight than a non-expert. In fact, weak evidence exists that for obscure questions with knowable answers (e.g., see Dalkey, 1970) experts, without use of reference books or data bases, are able to give more accurate answers then non-experts.

But how can experts be identified *a priori*? Dalkey and others have used various systems of self-rating; for example,

1. Are you an expert in this field, working in it daily?

2. Do you work in this field occasionally?

3. Are you knowledgeable about this field through occasional professional reading?

4. Would you classify yourself as an informed layman?

5. Are you uninformed about this field?

By assigning "points" for each level of self-rated expertise in the analysis of the group's response, answers can be discounted for lack of expertise to arrive at a group opinion.

If this approach were to be used in a Delphi study, then all questions would be sent to all participants. Each question would have appended a self-rating section. This approach would

provide a weighing yardstick but has several disadvantages.

- Experts for many of the questions posed may not exist;
- Even if experts exist, should their answers carry more weight than nonexperts? (After all, if experts could answer the questions associated with poverty, for example, why haven't they already done so?)
- People with good and useful suggestions could be dissuaded from contributing since they might have to admit a self-depreciating level of ignorance.

Lock-and-Key Approach. In this approach, administrators would attempt to match the capabilities of participants with the requirements of the questions. Each question and each respondent is "profiled," and questions are addressed only to those respondents whose profiles match.

Respondent profiles can come from a questionnaire that is completed by the respondent at the time of joining the panel and can be updated periodically. The profile would consist of best (or several) choice alternatives of the following sort:

1. What term best describes your occupation?

Scientist_____	Politician_____	Physician_____ etc.
Business person____	Cleric_____	Service worker_____
Engineer_____	Artist_____	Manufacturer_____
Teacher_____	Publisher_____	Retailer_____

2. If scientist, what is your discipline? (Parallel questions for engineers, business persons teachers, etc.)

Particle physics	Genetics	Biomedicine	etc.
Organic chemistry	Nuclear physics	Astronomy	
Materials	Psychology	Economics	
Social Science	Political science	Agriculture	

3. Would you consider yourself a:

Generalist
Specialist

4. Are you interested in (or do you have experience in):

Technological forecasting	Political processes	Utopias	etc.
Value-related issues	Policy analysis		
Science Policy	World affairs		

In other words, the respondents would be asked to identify themselves with respect to the same descriptors that will be used to profile the questions. Before any questionnaire is sent out, the

profile would be categorized according to the same set of descriptors that the respondents used to describe themselves. Then, through a matching process of the sort described below, the degree of match between a question and each respondent could be scored:

Desired Question Attributes	*Weight*	Resp No. 1	Resp No. 2	Resp No. 3	Resp No. 4	Resp No. 5
Scientist	5	0	1	1	0	0
Materials	10	0	1	0	1	0
Generalist	3	0	1	1	1	0
Science Policy	2	1	0	1	0	0
Techn Fcsting	1	1	0	1	0	0
SCORE		3	18	11	13	0

In this simulated example, suppose that the administrators felt that the question could best be answered by a materials scientist, who considered himself a generalist, with experience in science policy and technological forecasting. Not each of these descriptors was viewed as being equally important, however. Being a scientist was weighed at 5; being an expert in materials was weighed at 10; etc. Suppose further that data exist for each potential respondent with respect to each attribute (from the questionnaire completed at the time they joined the panel), so a simple scoring matrix could be constructed that assessed each respondent with respect to the desired attributes. By taking weighted sums, each respondent could be assigned a score. In this example, respondent 1 scored 3, and respondent 2 scored 18. After completing such an exercise, respondents could be rank ordered by score and the top set selected as a subpanel to which the question would be addressed. The number of people in a subpanel might be kept common across all questions, determined by budget considerations, or limited to say 90 percent of the top possible score.

Of course, if the number of respondents was large, this process would have to be automated. The program would have access to a data base of the respondents' characteristics. It would call for question attributes and weights, automatically scan the data base, score each respondent and present a different rank-ordered list of respondents for each question. The selected number of these could then be addressed automatically.

Free Text Search. This approach is similar to the lock-and-key approach except that it envisions a more open ended and fully automated means of matching question requirements to respondent expertise. Suppose that each respondent was asked to submit a free text resume to the project when they joined. The resumes might be constrained to a given length and follow a predetermined format requiring, for example, that the respondents describe their current and past activities, interests, publications, etc. A data base would be formed of these resumes.

Now, when each question is formulated, a set of keywords would be associated with the question, or the significant words in the question itself would be viewed as a keyword set. As before, each keyword would be assigned a weight. The data base of resumes would be searched for these keywords and, based on "hits," each potential respondent would receive a score. They

could be rank ordered as before.

Narrowing the Universe. In this approach, rather than asking about the respondents' expertise and interest, we let them select questions of interest and use their past selections to guide the routing of questions in future inquires. Suppose, after a respondent joined the network they received, for a time, all questions. They would, however, be instructed to answer only questions of interest to them and which they felt competent to address. Then, over time, the administrators could learn which categories of questions each respondent included. When sufficient data were available, only the types of questions of past interest would be sent to them.

The danger, of course, is that this approach is limiting. If the scanning process is cut off too early, then respondent/question matches that could have been productive might be missed. To minimize this potential detriment, respondents might be sent brief descriptions of other questions not included in their questionnaires and asked if any of these might be of interest to them.

Past Performance Scoring. Here we define an expert as a person who can provide correct answers to difficult questions. We want experts to answer questions about the future on the network; therefore, we will first pose qualifying questions that experts should get right. Answering these qualifying questions correctly is the "ticket of admission" to the real inquiry. For example, if the administrators wanted to make a forecast of the future value of the German mark, the qualifying question might be "On yesterday's date, what was the value of the German mark vis-à-vis the U.S. dollar?" Only those people who answered close to correctly would be presented the future-oriented question.

This approach has obvious problems. The test questions would have to be designed with meticulous care to be appropriate. More important, many respondents might feel that such a test of their competence is inappropriate.

A variation on this theme is to keep a record of each respondent's "batting average" by subject; that is, the ratio of correct to total forecasts made in the past. Important questions would be sent to those who have done well in the past. While initially attractive, this approach has a number of shortcomings also. First, it would have to be built up over time. Different respondents would have a different number of "times at bat." Finally, whether a forecast has indeed occurred is often difficult to tell, even in retrospect.

The Tree Approach. In this approach, a sequential series of questions is asked of the respondents, each requiring a more detailed knowledge of the field. At some point along the line, the respondents self-disqualify. Here's an example. Suppose our question has to do with the future of the electoral college system of elections in the United States: *Will the electoral college system of election in the United States change; if so, to what form and when?* The administrators might ask:

Domain Question: Have you ever thought much about how presidents are elected in the United States? (Yes, No)

If "no," no further questions.

Qualifying Question: Can you name three other ways in which this might be done? (Yes, No)

If "no," no further questions.

Actual Question: Do you think the electoral system in the United States will change within the next 20 years? (Yes, No)

If "no," then ask: Why not?
If "yes," then ask: Why? To what form? When?

The disadvantage of this approach is its complexity and its need for formulating precise questions. If this could be tolerated, it would certainly lead to appropriate matching of question to respondent.

The general principles of a Delphi study, anonymity, and feedback of prior round information to the current round, have been used in several novel ways in recent work by the Millennium Project. Some highlights of these applications are listed below to illustrate the wide range of possibilities for this method.

The question of rewards. One of the key principles of Delphi has been anonymity of respondents, at least with respect to "who said what." The successes of prediction markets (see the chapter on this method elsewhere on this CD ROM) raise the issue: would a built-in reward system improve accuracy? At one level this could mean allowing the respondents to specify that at least some of their specified contributions could be attributed to them. Or, as in the case of prediction markets, respondents might be given prizes or rewards if it is later found that they have produced accurate or otherwise useful forecasts. In prediction markets the rewards can be psychic when the currency of the market is "play" money or real when the contracts are bought and sold with real money or when "winners" are given prizes or donations to favorite charities are made in their name. At the very least the prediction contracts offer a means to encourage continued participation through the sequential Delphi rounds. Applied to Real Time Delphi, a reward system could encourage frequent revisits to the on line questionnaire.

Delphi has been used in scenario construction. Delphi was used in scenario construction in a study which dealt with **anti-terrorism**. The study began with a request for anti-terrorist scenarios distributed to listservs of the Millennium Project of the American Council for the United Nations University and the World Futures Studies Federation. The scenarios submitted in response to this request were analyzed to identify and rate policies and actions that might be useful in counterterrorism strategies. This work was posted on-line with a further request for comments, modifications to existing scenarios, and added scenarios. The submitted scenarios and others from outside this effort were reviewed to identify actions and policies that might be useful in counterterrorism strategies. The fifty-nine actions/policies identified were then submitted to an international panel for judgments about their effectiveness, plausibility, and unexpected potential downside risks. (Glenn and Gordon, 2002)

Another example of the use of the technique was made by the Millennium Project in forming

normative scenarios. Millennium Project participants identified and rated norms that were to be used as a normative scenario framework. In the first round, participants selected the following top four norms around which to form the scenario: environmental sustainability, plenty, global ethics (the identified and accepted), and peace. The others in order of preference were health, freedom, universal education access, equity, preservation of the human species, enlightenment, exciting and meaningful life, self-actualization, longevity, everyone has everything they want, and security. A normative scenario was devised around these norms; the scenario focused on actions to address the Global Challenges being tracked by the Project. These actions connected the present world to the normative future of 2050. For the second round, a scenario review panel of long-term normative-oriented participants was formed; they were asked to review and improve the draft of the scenario to illustrate optimistic possibilities for the future (Glenn and Gordon, 2002)

The Delphi approach has been used in construction of a **State of the Future Index (SOFI)**. In 1999–2000, the Global Lookout Panel of the Millennium Project was asked to identify indictors by which the status of 15 global challenges could be measured. These nominated indicators were subsequently evaluated by the panel in terms of their availability and usefulness. The results, plus a review of other index studies, were submitted to the Global Lookout Panel in 2001 to collect judgments about potential indicators for the SOFI. The respondents provided judgments about what the best (norm) and worst (dystopian) status was for the indicator in 2011. They also rated the importance of reaching the norm and dystopian states. The criteria for assigning a high weight to a variable were: the number of people affected; the significance of the effect; whether some groups seem to be affected differentially; the time over which the effect would be felt; and whether the effect is reversible. Millennium Project staff worked with the variables identified in this questionnaire, obtaining 20 years (where possible) of historical data from the most authoritative sources, and forecasting each variable using a time series approach (and later, using Trend Impact Analysis) to form a State of the Future forecast. (Glenn and Gordon, 2002

Delphi principles have been used in **policy studies**. Turoff wrote about this use of Delphi as early as 1970 (Turoff, 1970). As a recent example, in 2002-2003, the Millennium Project performed a study of routes to peace in the Middle East. In this application, the first questionnaire listed pre-conditions to peace and actions that could be taken to satisfy the preconditions. The panel was asked to add to the list of actions and to evaluate the set in terms of importance, likelihood of implementation, and their potential to evoke undesired or unintended downstream consequences. Extensive written commentary was generated by both the respondents and those who elected not to participate. The data were analyzed not only from the panel as a whole but also from the standpoint of the antagonists, using sub-samples of responses from Arab countries and Israel. From this body of data, the Project formed a set of normative scenarios that began from points of agreement and illustrated paths to peace. (Glenn and Gordon, 2003)

VI. SAMPLES OF APPLICATIONS

The Delphi method can be used in almost any forecasting application. . The Millennium Feasibility Project, performed by the United Nations University for the U.S. Environmental Protection Agency in early 1993, produced an application of potential interest. This study included a Delphi that examined important future events and policies that could affect world population growth and the environment in the next 25 years and will be used as an example here.

The respondents for this Delphi had self-identified themselves in a prior study as expert or greatly interested in population or in the environment; this list was augmented by recommendations by the study's sponsor (EPA), a literature scan, and recommendations of people who had already accepted our invitation to participate. In the first round, 76 questionnaires were sent out via airmail, fax, and e-mail; 42 responses were received. The second round went to the same list. Of those who completed the first round, 25 responded to the second round.

The questionnaires were first sent to a small planning committee, largely in open-ended form; the responses from the smaller group helped define a much more extensive format. The questions dealt with past history (e.g., "What developments and trends have been important in shaping today's population growth patterns?") and future developments (e.g., "Do you think these trends will continue? What other novel developments will deflect the trends in the future? What will be their consequences?").

The design of the questions was straightforward. The subjects of population and environment were dealt with in separate sections of the questionnaire, but the flow for each was largely the same. In the first round, questions focused on past forces for change, both from the standpoint of historical importance and future influence. The researchers provided the group with an initial list and asked them to extend the list and provide judgments about historical and future importance. They then asked about new forces for change and potential future unprecedented events; again, the researchers provided a list, asked for other ideas that might extend the list, and requested judgments about the likelihood and impact of the future developments. In the case of population, the questionnaire also called for direct estimates about the future size of the population and population growth rates in several countries and regions. It also asked about factors that might be different between developed and developing countries. In the case of the environment, the questionnaire included an additional section dealing with the interaction between population and the environment.

The second round fed back the newly suggested items to the panel and requested judgments similar to those called for in the first round. It also asked for judgments about policies that might be implemented to improve the future situation. These policy questions focused on the one or two future events that were judged to have low probability but high favorable impact, if they were to occur. This question asked, in effect, what might be done to improve the probabilities of these events.

To illustrate the nature of the questions asked, the population section of the second questionnaire

is reproduced below. The full questionnaires and the results of the study are included in reference (Gordon and Glenn, 1993).

UNU MILLENNIUM PROJECT FEASIBILITY STUDY DELPHI ON ENVIRONMENT AND POPULATION — ROUND 2 QUESTIONNAIRE

1. POPULATION SECTION

The first round of this Delphi on Population and Environment asked the panel to estimate population and growth rates for several regions to the year 2018. **Table 1** below lists the panel's median responses. The inter-quartile ranges for the population estimates are shown in parentheses next to the population estimates. (The inter-quartile range is from 25% above the lowest to 25% below the highest estimates; hence, 50% of the estimates fall within this range.) For purposes of comparison, 1992 year-end population and growth rates are listed first (cities are 1990 data).

Table 1. 1992 and 25 year forecasts of population from Population & Environment Delphi I

Region	1992 Population	Rates	Median Forecast (& Ranges)	Rates
World	5.4 billion	1.7%	8.0 billion (7.0 - 8.5)	1.5 %
Developed Countries	1.2 billion	0.5%	1.4 billion (1.3 - 1.4)	0.35%
Developing Countries	4.2 billion	2.0%	6.5 billion (5.7 - 7.0)	1.8 %
Africa	654 million	3.0%	1.2 billion (1.0 - 1.4)	2.5 %
China	1.2 billion	1.3%	1.5 billion (1.4 - 1.6)	1.0 %
India	883 million	2.0%	1.3 billion (1.2 - 1.4)	1.5 %
United States	256 million	0.8%	300 million (300 - 320)	0.65%
Brazil	151 million	1.9%	230 million (210 - 250)	1.6 %
Japan	124 million	0.3%	130 million (130 - 137)	0.2 %
Iran	60 million	3.3 %%	105 million (100 - 120)	3.0 % %%%% %%
Mexico	88 million	2.3% %	140 million (128 - 140)	2.0 %
Tokyo	18.1 million	0.9%	21 million (20 - 22)	0.5 %
Mexico City	20.2 million	2.9%	32.5 million (30 - 35)	2.0 %

In the first round, the panelists were asked to rate some forces that led to the reduction of the

world population growth rate from 2.06% in the late 1960s to the 1.7% then current and to assess how these forces might change over the next 25 years. **Table 2** below shows the average of the panel's judgments. In making these assessments, they used **Scale A** shown below.

Scale A

Historic Influence	***Future Influence***
1 = Very important	1 = Greatly increasing in importance
2 = Important	2 = Increasing in importance
3 = Marginally important	3 = Remaining the same in importance
4 = Unimportant	4 = Decreasing in importance
5 = Counter impact	5 = No longer a factor, or mixed

Table 2 - The importance of some historic factors (first number) on global population growth and possible future changes in importance (second number)

- 2.0 2.2 Availability of inexpensive, simple effective contraceptives;
- 2.0 2.1 Family planning and public health programs
- 2.0 2.8 China's population policy
- 2.1 2.1 Government policies in developing countries that encourage smaller families
- 2.1 2.0 Increasing number of years that women attend school
- 2.1 2.3 Rising incomes and the spread of middle class values
- 2.5 2.4 Rise of "woman's power."

QUESTION 1.1 We asked the panel to suggest additional forces that might be responsible for the historic changes in population growth. Many new suggestions were received. Respondents were asked to review the list presented below and to provide their judgments about the historical and possible future importance of these forces using **Scale A,** both historically **and** over the next 25 years..

1.1.1 ___ ___ Move away from agricultural society/primary sector

1.1.2 ___ ___ Demonstration by the North that fewer children can mean more wealth

1.1.3 ___ ___ Legitimization of contraception

1.1.4 ___ ___ Availability of male contraceptives

1.1.5 ___ ___ Decreasing Catholic Church's social influence

1.1.6 ___ ___ Increasing futuristic orientation

1.1.7 ___ ___ War, famine, disease, and pestilence

1.1.8 ___ ___ Environmental deterioration

1.1.9 ___ ___ Spread of new communications media (television, etc)

1.1.10 ___ __ Education about relation of environment and population

1.1.11 ___ __ Decreases in infant and child mortality

1.1.12 ___ ___ Improved literacy by improved children's schooling

Population growth rates have remained high in developing countries. Using the **Scale A** above, they were asked to judge the importance of historical influences and how they might change in importance over the next 25 years. **Table 3** presents the panel's estimates. Historical assessment is the first number; future assessment is the second number.

Table 3 - Some reasons for high population growth in developing countries

1.6 3.0 Need of children for social security, to support parents in their old age.
1.8 2.4 Poverty
2.0 2.7 Low levels of literacy
2.2 2.8 Lack of information and access to contraceptives
2.2 3.1 High infant mortality
2.8 3.4 Government policies supporting large families

QUESTION 1.2 In the first round, the panelists were also asked to suggest other forces that might account for population growth rates remaining high in many developing countries. Many new suggestions were received. They were asked to review the developments and to provide judgments about their historical and possible future importance using **Scale A.**

1.2.1 ___ ___ Masculinity associated with increasing numbers of children

1.2.2 ___ ___ Close adherence to religious tenets that lead to avoidance of contraceptives

1.1.3 ___ ___ Low levels of literacy and lack of understanding ecological view of planet Earth

1.2.4 ___ ___ Discrimination against women (little autonomy or education, and lack of social power)

1.2.5 ___ ___ Belief by governments that larger populations mean greater political strength

1.2.6 ___ ___ Rural areas receive less attention from population programs than urban areas

1.2.7 ___ ___ Family-based, labor-intensive economies needing children's input

1.2.8 ___ ___ Liberal immigration policies in richer countries

In the first round, the panelists were asked to assess new forces and unprecedented events that might influence population growth in the future. They were asked for judgments about the likelihood of occurrence and impacts over the next 25 years. **Table 4** shows the average of the

panel's responses about the events included in the first round. In making these assessments, they used **Scale B** shown below.

Scale B	***Likelihood of occurrence***	***Eventual impact on population within the next 25 years of growth***
	1 = almost certain	1 = reduces growth rate by 30% or more
	2 = likely	2 = reduces growth rate by 5-30%.
	3 = even or 50/50 chance	3 = no impact.
	4 = unlikely	4 = increases growth rate by 5-30%.
	5 = almost impossible	5 = increases growth rate by 30% or more

Table 4 Likelihood of occurrence of new forces or unprecedented events (first number) that might influence population growth and eventual impact over the next 25 years (second number)

1.8 2.7 Simple test for identifying the sex of unborn children
1.9 3.4 Increasing survival in middle age and early old age due to curing or improved therapy for heart disease, cancer and stroke
2.1 2.3 Simple, safe, effective male birth control pill
2.3 2.4 Changes in death rates due to spread of AIDS in developing countries
2.4 2.6 Massive starvation of the scale of Somalia at least once every three years
2.5 2.0 Long term (at least one year) contraceptives as widely used and accepted as birth control pills are today
2.6 2.4 New deadly viruses, including AIDS mutations
2.8 2.6 Simple method for selecting sex at conception
2.9 2.1 Rising incomes in most developing countries
3.3 2.3 Change in the Vatican's position: use of contraceptives becomes available without limit
3.4 2.1 Two children per family becomes the social norm in the majority of the developing world

QUESTION 1.3 In the first round, the panelists were asked to suggest other new forces or unprecedented events that could influence population growth. Many new suggestions were received. They were asked to review the list below and provide their judgments about their likelihood and impacts, if they were to occur over the next 25 years using **Scale B** in the first space in front of each item for the likelihood of occurrence and in the second space for the item's future impact over the next 25 years.

1.3.1 ___ ___ Widespread use of relatively cheap and easy ways to affect the aging process, resulting in diminished mortality and extension of the life span by about 5 years

1.3.2 ___ ___ Increasing sterility, worldwide, by 10% due to environmental degradation, higher level of stress, and other factors

1.3.3 ___ ___ Increasing impotency by 10% due to environmental degradation, higher level of stress, and other factors

1.3.4 ___ ___ Public health programs decrease mortality of infant and young adults by 5-10%

1.3.5 ___ ___ 3% of births via new methods of impregnation and prenatal development ("Outside womb" fertility, artificial insemination, surrogate motherhood, other such techniques)

1.3.6 ___ ___ Important negative changes in the environment (e.g., accumulation of toxic wastes, failures of mono-agriculture crops, contamination of drinking water) resulting in increased mortality

1.3.7 ___ ___ Doubling, worldwide, of today's level of male and female homosexuality

1.3.8 ___ ___ Basic economic needs met for 90% of global population (minimum acceptable health care, food, and shelter)

1.3.9 ___ ___ Successful new prototype habitats in oceans, cold regions, or in earth orbit stimulates popular frontier spirit and alternatives to previous urbanization patterns

1.3.10___ ___ 20% increase of people active in religions that encourage higher fertility

QUESTION 1.4 With the information provided by the panel in round I, we identified two possible future events that were judged to be unlikely, but nevertheless potentially effective in reducing population growth rates, if they were to occur. Both have been considered before and one is the focus of large international programs. Without repeating what has been said and tried many times, we ask below for your suggestions about novel policy approaches that you think might be practical, and, if implemented, improve the probability of these developments. Please write your suggestions in the space provided below. Change the statements if you wish. Add other policy domains if you wish. Please be as specific as possible and use separate sheets of paper or electronic space as you need.

1.4.1 Novel policies that could lead to a social norm of two children per family throughout the majority of the developing world:

1.4.2 Novel policies that could lead to the Vatican's acceptance of contraception without limit:

1.4.3. Other policy areas of your selection and how you would address it:

To illustrate the quality of the responses received, some of the results of this questionnaire are summarized below; a complete analysis- including summaries of answers to all questions- appears in Gordon and Glenn, 1993.

Using the scale for historical influence, the panel found the following ten developments most significant in the evolution of world population growth.

The Ten Most Important Historical Factors

2.0 Availability of inexpensive, simple effective contraceptives;
2.0 Family planning and public health programs
2.0 China's population policy
2.0 Legitimization of contraception
2.0 Decreases in infant and child mortality
2.1 Government policies in developing countries that encourage smaller families
2.1 Increasing number of years that women attend school
2.1 Rising incomes and the spread of middle class values
2.2 Move away from agricultural society/primary sector
2.3 Spread of new communications media (television, etc.)

Further, the panel judged, using the scale for future influence, that among forces already in play, the following would be the ten with the most greatly increasing importance over the next 25 years

The Ten Historical Forces That Promise to be of Increasing Importance
1.8 Spread of new communications media (Television, etc.)
2.0 Increasing number of years that women attend school
2.0 Environmental deterioration
2.1 Government policies in developing countries that encourage smaller families
2.1 Family planning and public health programs
2.1 Improved literacy by improved children's schooling
2.2 Availability of inexpensive, simple effective contraceptives
2.2 Legitimization of contraception
2.2 Education about relation of environment and population
2.2 Decreases in infant and child mortality

The questionnaire included similar questions about developing countries: reasons for high population growth in these countries and the potential importance of these forces over the next 25 years. Here is a summary of the top items:

The Ten Most Important Reasons for High Population Growth in Developing Countries
1.6 Need of children for social security, to support parents in their old age.
1.7 Discrimination against women (little autonomy, education, and lack of social power)

1.8 Poverty
1.8 Low levels of literacy and lack of understanding an ecological view of planet Earth
2.0 Family-based, labor-intensive economies needing children's input
2.1 Close adherence to religious tenets that lead to avoidance of contraceptives
2.1 Rural areas receive less attention from population programs than urban areas
2.2 Lack of information and access to contraceptives
2.2 High infant mortality
2.3 Masculinity associated with increasing numbers of children

The Ten Historical Forces That Promise to be of Increasing Importance in Developing Countries over the Next 25 Years
2.2 Low levels of literacy and lack of understanding ecological view of planet Earth
2.4 Poverty
2.6 Rural areas receive less attention from population programs than urban areas
2.7 Low levels of literacy
2.8 Discrimination against women (little autonomy or education, and lack of social power)
2. 9 Lack of information and access to contraceptives
3.0 Need of children for social security, to support parents in their old age.
3.1 Family-based, labor-intensive economies needing children's input
3.1 High infant mortality
3.2 Belief by governments that larger populations mean greater political strength

With the information provided by the panel in round I, we also identified two possible future events that were judged to be unlikely, but nevertheless potentially effective in reducing population growth rates, if they were to occur. These were;

Impact	**Probability**	**Development**
2.1	3.4	Two children per family becomes the social norm in the majority of the developing world.
2.3	3.3	Change in the Vatican's position; use of contraceptives becomes available without limit.

These items were of particular interest to us since they represented potential policy opportunities. Both have been considered before and one is the focus of large international programs. We asked for suggestions about novel policy approaches that might be practical, and if implemented, improve the probability of these developments. We also asked our respondents to not repeat, if possible, previously made suggestions. Some of the panel's ideas follow:

Novel policies that could lead to a social norm of two children per family throughout the majority of the developing world:

- Family-size taxes that increase substantially for every child over two.

- Prohibition of child labor.

- Policy focus on women's access to work and education.

- Free circulation of contraceptives in public health programs supported by churches and other religious organizations.

- International information utilities in education, health, and training in information society.

- With enough effort and resources, the present policies (national, bilateral, and multi-lateral) that try to make family planning services universally available and to promote information, education and communication about family planning, should make the two-child norm virtually universal within 25 years.

- Without significant reductions in poverty and infant mortality, and increase in women's education and empowerment, these policies would need to be coercive in nature or at least provide strong economic disincentives to having many children.

Some Novel Policies That Could Lead to the Vatican's Acceptance of Contraception Without Limit:

- Remove Holy See from the United Nations on grounds that it is not really a country and give it the same status as the World Council of Churches.

- Promotion of contraceptives accompanied with strong dissemination of moral values.

- Theological doctrine developed by US. Catholic Bishops in support of sustainable development (at UNCED 1992).

- Policies that focus on responsibilities of women to make choices.

- Only decline of organized religion is likely to affect the situation.

- Allow priests and nuns to marry and pay for the raising of children.

- Reconsideration of the theology of St. Thomas Aquinas - especially the discarding of the Thomastic view of "natural law" as it applied to human sexuality. Christ said nothing whatsoever about human sexuality. Separate the notion of procreation as the only "natural end" from that of enjoyment.

We also asked respondents to suggest other areas that were ripe for policy intervention. Here are some they suggested:

- Research for long-term implantable ovulation suppression device.

- Global televised debate on population policy, environmental protection, and social ethics.

- Sex education in schools

- All nations will have to eventually adopt policies that clearly state the freedom of individual choice.

- Policies that emphasizes people to take charge of their lives and reduce dependence on governments.

BIBLIOGRAPHY

Amant, R., "Comparison of Delphi Forecasting Studies in 1964 and 1969," *Futures*, Vol 2, No. 1, March 1970. This article is an early attempt to determine just how accurate earlier forecasts made by the Delphi method have proven to be. It showed that scientific and technological forecasts were indeed likely to be more accurate than forecasts in softer areas, such as fashion, politics, or social behavior. It also showed that sometimes, whether an event, even one succinctly stated, had indeed occurred was quite difficult to tell.

Azani, Hossein and Khorramshahgol, Reza, "Analytic Delphi Method," *Engineering Costs and Production Economics*, July 1990.

Biji, Rob, "Delphi in a Future Scenario Study on Mental Health and Mental Health Care," *Futures*, April 1992.

Dalkey, Norman, B. Brown and S.W. Cochran, "The Delphi Method III: Use of Self Ratings to Improve Group Estimates," *Technological Forecasting and Social Change*, 1, 1970

Dionne, C.E., et al., "A consensus approach toward the standardization of back pain definitions for use in prevalence studies, *Spine*, Volume 33, Issue 1, January 2008, Pages 95-103

Gibson, Lay James and Miller, Mark, "A Delphi Model for Planning 'Preemptive' Regional Economic Diversification," *Economic Development Review*, Spring 1990.

Ginsburg, Alan, "Integrating Evaluation Into Decisionmaking," *Public Manager*, Winter 1992-93.

Goldfisher, Ken, "Modified Delphi: A Concept for Product Forecasting," *Journal of Business Forecasting*, Winter 1992-93.

Glenn, Jerome and Theodore Gordon, *State of the Future 2002*, Millennium Project, American Council for the United Nations University, Washington, DC. 2002.

Glenn, Jerome and Theodore Gordon, *State of the Future 2003*, Millennium Project, American Council for the United Nations University, Washington DC. 2003.

Gordon, T. J. and Glenn, J. C., *Issues in Creating the Millennium Project: Initial Report from the Millennium Project Feasibility Study*, United Nations University, August 1993. This work was, in part, an exploration of the feasibility of conducting worldwide Delphis on issues of global importance, using scholars, futurists, and other interested and informed participants in panels that communicated by e-mail, fax, and air mail.

Gordon, T. J. and Helmer, Olaf, *Report on a Long Range Forecasting Study, RAND Corp.*, R-2982, 1964. This report is still interesting to read. It contains many innovations that are used in the analysis and presentation of Delphi results. For example, the reports presents arguments for

using median rather that means of the group's responses, shows how ranges of opinions can be presented graphically, distinguishes between likelihood and probability, etc.

Helmer, Olaf and Rescher, Nicholas, "On the Epistemology of the Inexact Sciences," *Management Sciences*, Vol. 6, No.1 (1959). In this landmark work Helmer and Rescher set out the philosophical backdrop for Delphi and set limits of expectation about what can and cannot be known when the questions being addressed fall into the category of "inexact science." They include all of social science and political science as "inexact."

Kastien, M. R., et al., "Delphi, The Issue of Reliability," *Technological Forecasting and Social Change*, Nov. 1993.

Lindstone, H., and Turoff, M. (ed.), *The Delphi Method*, Addison Wesley Publishing Co., 1975. One of the best detailed discussions of the Delphi method.

McCarthy, Kevin, "Comment on the 'Analytic Delphi Method'," *International Journal of Production Economics*, May 1992.

Mitchell, Vincent-Wayne, "Using Delphi to Forecast New Technology Industries," *Marketing Intelligence and Planning*, Vol 10, Issue 2, 1992.

Nimgade, Ashok and Sonk, Joseph, "Ordinal Quantification of the Concept of Technology Through Pooled Consensus," *R&D Management*, Jan. 1991.

Olshfski, Dorothy, and Joseph, Alma, "Assessing Training Needs of Executives Using the Delphi Technique," *Public Productivity and Management Review*, Spring 1991.

Pare, G, C.Sicotte, M. Jaana, and D. Girouard, "Prioritizing clinical information system project risk factors: A Delphi study", Proceedings of the Annual Hawaii International Conference on System Sciences, 2008.

Rohde, William, "Past Predictions Take Shape," *Healthcare Executive*, May/June 1991.

Rowe, Gene, George Wright, "Expert Opinions in Forecasting: The Role of the Delphi Technique," a chapter in *Principles of Forecasting,* J. Scott Armstrong (ed) Springer (2001).

Rowe, Gene, George Wright, and F. Bolger, "Delphi: A Revaluation of Research and Theory," *Technological Forecasting and Social Change*, Vol. 39, May 1991.

Russo, David M., and McLaughlin, "The Year 2000: A Food Industry Forecast," *Agribusiness, November* 1992.

Shota, Ushio, "The Future of High Tech: Forecasts for the Next Decade," *Tokyo Business Today*, April 1993.

Technological Forecasting and Social Change, Special Issue on Delphi, 1975, no. 2. Discusses critiques of the method.

Turoff, Murray "The Design of a Policy Delphi," *Technological Forecasting and Social Change 2*, No. 2 (1970).

Turoff, Murray, Delphi Conferencing: Computer-Based Conferencing with Anonymity, Technological Forecasting and Social Change 3, 159-204, 1972.

Turoff,Murray, Starr Roxanne Hiltz, Hee-Kyung Cho, Zheng Li, and Yuanqiong Wang, Social Decision Support Systems (SDSS), Proceedings of the 35th Hawaii International Conference on System Sciences, 2002

Tversky, A. and D. Kahneman, "The Framing of Decisions and the Psychology of Choice," Science, 211, 453-458, January 30, 1981.
http://psych.hanover.edu/classes/Cognition/papers/tversky81.pdf

Vickers, Brent, "Using GDSS to Examine the Future European Automobile Market," *Futures*, October 1992.

Woudenberg, Fred, "An Evaluation of Delphi," *Technological Forecasting and Social Change*, Sept. 1991. Contains a broad review of the literature on Delphi and many references to the method and past studies.