ENCYCLOPEDIA OF
MEASUREMENT
AND
STATISTICS

VOLUME 1

EDITED BY
NEIL J. SALKIND

ENCYCLOPEDIA OF
MEASUREMENT
AND
STATISTICS

VOLUME 1

EDITOR
NEIL J. SALKIND
UNIVERSITY OF KANSAS

MANAGING EDITOR
KRISTIN RASMUSSEN
UNIVERSITY OF KANSAS

Contributors

Francisco J. Abad
Universidad Autonoma de Madrid

Inmaculada Aban
University of Alabama at Birmingham

Hervé Abdi
University of Texas at Dallas

Phillip L. Ackerman
Georgia Institute of Technology

Demetrios S. Alexopoulos
University of Patras, Greece

Audrey Amrein-Beardsley
Arizona State University

Lauren E. Auld
DePauw University

Carrie R. Ball
University of Wisconsin–Madison

Kimberly A. Barchard
University of Nevada, Las Vegas

Jonathan Barzilai
Dalhousie University

Edward J. Bedrick
University of New Mexico

Mark L. Berenson
Montclair State University

Dongsheng Bi
University of Nebraska Lincoln

Damian P. Birney
University of Sydney

David M. Boynton
Saint Michael's College

Bruce A. Bracken
College of William & Mary

Jennifer Bragger
Montclair State University

Gary G. Brannigan
State University of New York–Plattsburgh

Ernest W. Brewer
University of Tennessee

Carolyn Brodbeck
Chapman University

Sarah Brookhart
American Psychological Society

Duane Brown
University of North Carolina, Chapel Hill

Jennifer Ann Brown
University of Canterbury

Shawn T. Bubany
University of Minnesota

However, $\bar{X}_n$ is the limit of a sequence of Bayes estimators. More precisely, consider as a prior distribution on Θ the normal distribution with mean m and variance d^2. That is, the prior density is given as follows:

$$\lambda(\theta) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2d^2}(\theta - \mu)^2 \right\}.$$

The Bayes estimator with respect to this prior and squared error loss is

$$a(X_1, \ldots, X_n) = \left\{ \frac{(n/\sigma^2)}{n/\sigma^2 + 1/d^2} \right\} \bar{X}_n + \left\{ \frac{(1/d^2)}{n/\sigma^2 + 1/d^2} \right\} \mu.$$

Thus, the Bayes estimator is a convex combination of the sample mean $\bar{X}_n$ and m, the mean of the prior distribution. Letting $b \to \infty$, it is seen that $\bar{X}_n$ arises as a limit of Bayes estimators.

—*Javier Rojo*

See also Bayesian Statistics; Evidence-Based Practice

Further Reading

Bergus, G. R., & Hamm, R. (1995). Clinical practice: How physicians make medical decisions and why medical decision making can help. *Primary Care, 22*(2), 167–180.

Braddock, C. H., III, Edwards, K. A., Hasenberg, N. M., Laidley, T. L., & Levinson, W. (1999). Informed decision making in outpatient practice: Time to get back to basics. *Journal of the American Medical Association, 282*(24), 2313–2320.

Green, D., & Shapiro, I. (1994). *Pathologies of rational choice theory: A critique of applications in political science.* New Haven, CT: Yale University Press.

Lehmann, E. L. (1983). *Theory of point estimation.* New York: Wiley.

Lehmann, E. L. (1991). *Testing statistical hypotheses.* Belmont, CA: Brooks/Cole.

Lehmann, E. L. (2004). Optimality and symposia: Some history. In J. Rojo & V. Perez-Abreu (Eds.), *The First Erich L. Lehmann Symposium: Optimality.* IMS Lecture Notes and Monograph Series, Vol. 44.

Neyman, J., & Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society, Series A, 231,* 289–337.

Rojo, J. (1987). On the admissibility of c $\bar{X}$ + d with respect to the linex loss function. *Commun. Statist. Theory and Meth., 16,* 3745–3748.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology, 28,* 1–39.

Varian, H. R. (1975). A Bayesian approach to real estate assessment. In S. E. Fienberg & A. Zellner (Eds.), *Studies in Bayesian econometrics and statistics in honor of Leonard J. Savage* (pp. 195–208). Amsterdam: Elsevier North-Holland.

Wald, A. (1950). *Statistical decision functions.* New York: Wiley.

Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association, 81,* 446–451.

Delphi Technique

The Delphi technique is a means of collecting data from a diverse group of people for the purpose of reaching a consensus. This entry presents the basic process of the Delphi technique as well as some variations on the process that can be used to meet specific needs. Although the Delphi technique allows for refinement of original ideas and therefore promotes high-quality decisions, it can be time consuming and subject to bias. Examples of how the process can be used in research are provided. From this entry, readers can determine if the Delphi technique is appropriate for their particular situations.

According to S. J. Adams, the Delphi technique provides a representation of varied backgrounds, and it prevents individuals with strong personalities from dominating a group. The purpose is to obtain information from participants to help in the areas of problem solving, planning, and decision making. The Delphi technique is a way to reach a consensus among a group of experts.

The RAND Corporation developed the technique during the 1950s as an approach to forecasting the likelihood and the potential impact of Russian bombing attacks on the United States. The approach was named for the Oracle of Delphi of Greek mythology. It was soon adopted by technological forecasting experts and eventually found its way into other types of research.

Theoretical Basis for the Delphi Technique

Consensus Theory

According to a variety of researchers, the objective of users of the Delphi technique is to achieve consensus. Some proponents of consensus theory believe that building consensus offers opportunity for communal renewal and for achieving group commitment to common goals.

Anonymity

Some researchers and theorists believe that anonymity is helpful for generating quality ideas. Others expect that using the Delphi process discourages individual dominance and simultaneously encourages each person to share his or her ideas without fear of intimidation.

Divergent Thought

Some researchers observe that divergent thinking occurs when individuals or groups are introduced to minority opinions. Anonymity and exposure to a variety of viewpoints contributes to improved creativity and decision making.

Purpose and Uses in Research

Researchers have used the Delphi technique for gathering broad-based opinions from experts, refining their views, and reaching consensus on predictions and plans for dealing with complex issues. The data generated have been used in forecasting, public budgeting, and goal setting. Decision makers in such diverse disciplines as education, safety management, family therapy research, environmental studies, government, medicine, and community health have relied on Delphi for all or portions of their research data.

Delphi Technique Process

Delphi technique involved several carefully structured steps. It bore some resemblance to Nominal Group Technique (NGT) in that with both processes,

individual contributions were made anonymously. However, the standard format for Delphi did not require participants to meet. Thus, not only were responses anonymous, but even the identity of other participants might be unknown to the group. The procedure involved two to four rounds of responses. However, prior to the first round, primary stakeholders had to do the following:

1. *Select a monitor or monitor group.* This person or persons should be experts both on the topic and on written communication skills.

2. *Select participants.* Participants usually were stakeholders as well. However, they could be noninvolved experts.

3. *Invite participants.* Selected participants were invited by telephone, mail, or e-mail to take part in the process.

4. *Develop a broad question or statement for consideration.* The monitor developed the initial question or statement, perhaps in conjunction with other stakeholders.

To begin the rounds, the monitor was responsible for (a) identifying and orienting participants; (b) getting the question to each participant; (c) receiving input from each participant; (d) summarizing the information; (e) sending the summary and a new, more focused question to the participants; and (f) determining that no more rounds were needed. The process concluded with a resolution. When consensus was reached, the resolution was announced to participants. Panel participants committed to the decision (see Figure 1).

It was recommended that 12 to 15 panel members were an appropriate size. Panel sizes ranged from a few to hundreds of members, depending on the research topic. A response rate of 70% or greater was typically acceptable. It was common for the iteration process to last only two or three rounds before consensus was reached.

Statistical Measures of Agreement

Panels commonly have used Likert scales to assess the rating of items. The Delphi monitor calculated
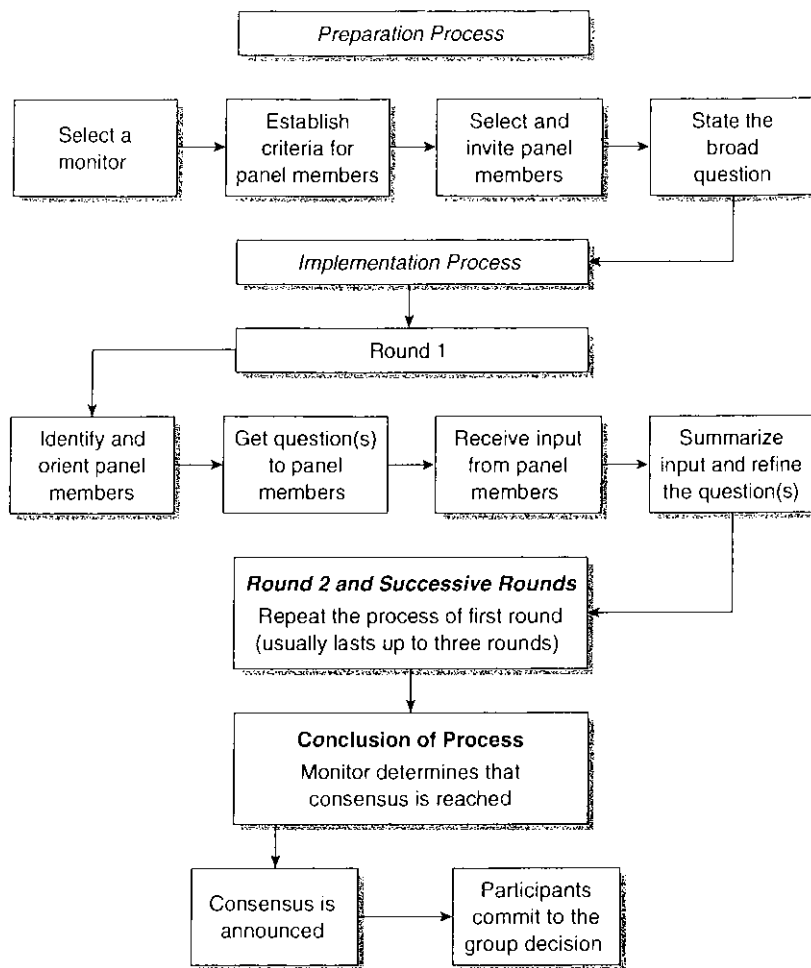
Figure 1   Flowchart of Delphi Technique

Variations of the Delphi Technique

Real-Time and Policy Delphi

Clayton identified three types of Delphi: (a) conventional, (b) real time, and (c) policy. The conventional process was presented earlier. Real-time Delphi differed in that it occurred face-to-face, within the context of a meeting or conference. Policy Delphi asked participants for information on which a decision was to be made. Policy Delphi and real-time Delphi could be combined. Consensus was not an objective in either case.

Combinations of Delphi With Nominal Group Technique

The possibility of voting when consensus could not be reached could be considered a variation on Delphi. It was a compromise of the ideal Delphi and borrowed the last phase from NGT.

summary statistics, such as the median, following each round and reported them to the panel members for consideration during the next round. Researchers found the median to represent the most common value provided by a panel member and cited the interquartile range (the middle half of the scores) as a measure of consensus. The smaller the interquartile range, the greater the consensus. Several studies supported using the median and the interquartile range as measures of agreement and consensus.

Delphi scholars reported means to panel members in successive rounds and standard deviations as measures of consensus. However, other researchers suggested that providing standard deviations to expert panels was misleading because they were not applicable to ordinal data.

Mixed Methods Studies

Some researchers have used mixed methodologies to complement one another. For example, one might use Delphi to determine items for a questionnaire that would be used in a mail survey regarding environmental policy.

Historic Delphi

In this variation, Strauss and Zeigler found that participants attempted to apply systematically the writings of classical political philosophers to current issues. That interesting study is presented in the *Historic Delphi Approach* section of this entry.

Reactive Delphi

A popular variation of Delphi was the reactive method, in which panel members reacted to pregenerated items or questions in Round 1, rather than producing a list of ideas. In this variation, the researcher prepared a list of items from a review of related literature, and the subject matter experts were asked to rate the importance of each item on some scale. The researchers found that such structured first rounds diminished the assessment properties found in the classic Delphi technique.

Advantages and Disadvantages of the Delphi Technique

Advantages of Delphi Technique

Effective structure. The process allowed participants to refine their original ideas. That resulted in high-quality decisions on complex issues. These results came from professionals who gained insights from one another's input during the successive rounds.

Fiscal economy. Little financial cost was involved in using the conventional Delphi technique. There are no travel costs, no need for accommodations, and usually no stipends for participants.

Collaboration. Several researchers noted that in reaching consensus, the Delphi technique fostered collaboration among parties who would be needed to carry out the group's decisions.

Disadvantages and Limitations of the Delphi Technique

Time requirements. Two aspects of time apply to Delphi. First, there is the requirement of the monitor(s) and of each participant. Also, the transmission of ideas could result in an overall time frame of several weeks. Difficulty in retaining participants throughout the process may be a problem. Sometimes, the latter could be lessened by using e-mail. However, that would require special care to maintain anonymity.

Inadequacy as sole method. According to several studies, when used alone, the Delphi technique was inadequate for forecasting. Also in this regard, one must consider the built-in dangers of bias. A discussion of such dangers follows.

Central tendency. Michigan State University Extension found that consensus building generally has involved finding a middle-of-the-road position, eliminating the extreme ends of the spectrum. This feature has caused some groups to feel that their views were rejected and that the process was rigged.

Bias. It is important to ensure that experts are not influenced by the researcher's objective. To guard against this, Delphi experiments usually use two or more separate groups of experts. There are numerous opportunities for introducing bias into the Delphi process, including (a) setting eligibility standards and soliciting participants, (b) formulating the questions, (c) summarizing participants' contributions, (d) rephrasing questions for successive rounds, (e) determining the number of rounds, (f) phrasing consensus statements, and (g) fostering commitment to decisions. For each step in the process, moderators must be accountable for keeping the process bias free.

Communication difficulties. Strauss and Zeigler discussed the possibility of misunderstandings, noting that the respondents may misunderstand the brief written inputs of the panelists. Others noted that Delphi relied heavily on the written communication skills of experts. This made the selection of participants especially difficult, because expertise in the field did not necessarily include communication expertise.

Ethical standards and need for trust. Conflict could arise from the many opportunities for bias using Delphi. Root causes could be intended or unintended bias or lack of group trust in the process. Without mutual trust, it would be impossible to reach consensus. Three types of disagreement involving ethics and trust are especially noteworthy:

1. *Forecast versus foresight.* Forecasting with the Delphi technique is used to predict what is likely to occur, whereas foresight implies that the process is guided toward a predetermined outcome. As mentioned elsewhere, the Delphi process cuts off extreme views and seeks a middle ground. Unfortunately, those whose opinions have been sacrificed may feel disenfranchised. Thus, when a facilitator records a

group's comments, the final outcome can be highly questionable.

2. *Consensus versus coalition.* The same principle applies here as to the forecasting versus foresight outcomes.

3. *Consensus and morality.* Among populations that must interrelate, pluralism was the only viable option. However, if Delphi technique participants represented a full spectrum of ethical values, absolutism, pluralism, and relativism, some participants might be open, some might be reluctantly persuaded, and some might feel excluded. If so, consensus could not be reached. However, the Delphi technique was designed especially to reach consensus on complex issues, and complex questions almost always involve moral values.

In a discussion of moral consensus, the following questions should be considered: Does consensus carry any moral authority? and Can groups ever achieve a valid consensus on issues of bioethics? Their considerations have raised questions about both the practicality and the propriety of using the Delphi technique to address the very issues for which the technique was designed. Perhaps one should consider the words of Mohandas Gandhi, "In matters of conscience, the law of majority has no place."

Research Applications

In this section is a brief overview of some of the institutional research applications using Delphi. They range from employee issues, such as selecting job candidates and handling occupational stress, to forecasting training needs and needed changes in human resource practices, to needs assessment.

Frazer and Sechrist examined the effects of occupational stress on employees in nuclear medicine, radiologic technology, and medical technology. They used the Delphi technique to determine 35 job stressors for each discipline. Improved communication strategies and managerial development were noted as solutions to occupational stress.

Olmstead-Schafer, Story, and Haughton used the Delphi method to forecast training needs of public health nutritionists. It was the consensus of their panel that communication, policy development, and managerial skills be included in the curriculum for training nutrition professionals.

Japanese firms used the Delphi method in forecasting needed changes in human resource practices. The panel made predictions regarding the year in which strongly held Japanese institutions of lifelong employment, seniority-focused compensation, and promotion from within at the exclusion of external recruiting would transition to practices consistent with Western cultures. The overall consensus was that it would take two decades to see significant changes.

Tavana, Kennedy, and Joglekar studied the effectiveness of the Delphi approach for ranking job candidates for a nursing management position. After two rounds, the experts reached consensus on the top applicant from a field of seven. Schuler found that the Delphi approach was beneficial in emergent and less structured subject areas such as human resource planning.

Finally, program evaluation was noted as another area for using Delphi. The Delphi technique is particularly useful for studies requiring a needs assessment.

Historic Delphi Application

Strauss and Zeigler conducted an interesting historic Delphi study. Their objective was to systematically scrutinize the great political philosophers of the past and to apply their wisdom to contemporary problems. Plato, Aristotle, Hobbes, Machiavelli, Swift, Burke, Rousseau, Locke, Marx, and Freud were the philosophers. Ten panels of six experts each (mostly university professors) represented the philosophers. The questionnaire contained 42 problem statements regarding serious issues in Western society, and each statement had a three-part question:

1. In general, what was political philosopher X's view on problem statement Y?

2. Based on your knowledge of political philosopher X, how would he have reacted to the problem statement in his own time?

3. If alive today, how would political philosopher X resolve the problem?

The second round consisted of multiple-choice options. The experts responded on a 5-point Likert scale that asked to what extent they agreed or disagreed with each item. The items in the second round were taken from each group's first-round responses.

The product of this academic exercise was a series of options for handling a variety of social problems based on Western philosophical thought. Strauss and Zeigler hoped that, in addition to accomplishing this pragmatic objective, their development of the historic Delphi approach would be a meaningful way for students to study philosophy.

Comparison of Delphi, Nominal Group, and Q-Sort Techniques

Delphi and NGT have many similarities. Each encourages divergent thought, preserves anonymity of participants' contributions, and is aimed at consensus. Each can be a powerful research technique for solving complex problems, and each has been adapted to a variety of needs through variants on the classical processes. Both processes require significant time commitments, and both are subject to bias. Both tend to discredit extreme positions and could alienate those stakeholders.

Q-Sort, on the other hand, is used primarily as an individual technique for developing theory related to human behavior and for identifying and describing human phenomena. The Q-Sort is a time-consuming process, as are Delphi and NGT. In contrast to those methods, Q-Sort researchers develop an instrument first, through literature review. The instrument is designed to measure using forced-choice options. Data collection is usually accomplished one-on-one. Table 1 depicts similarities and differences between these three research methods.

Summary

The Delphi technique was designed to identify the best solutions to complex organizational and other social problems; and researchers in diverse fields have used it in its conventional form and with several variants. However, the process is fraught with opportunities for contamination through bias, either actual or perceived. Necessary as it is in a pluralistic society, both the possibility and the propriety of reaching consensus remains illusive.

—*Ernest W. Brewer*

See also Decision Theory

Further Reading

Adams, S. J. (2001, October). Projecting the next decade in safety management: A Delphi technique study. *American Society of Safety Engineers, 32*, 26–29.

Table 1   Comparisons and Contrasts: Q-Sort, Delphi, and Nominal Group Technique

Name	Purpose	Data Collection	Primary Uses	Advantage	Disadvantage
Delphi	Consensus building	Group; anonymous	Medicine; social sciences	Divergent thinking; does not require panel participants to meet	Possible manipulation
NGT	Decision making	Group; anonymous	Social sciences	Divergent thinking	Possible manipulation; requires participants to meet
Q-Sort	Theory building; description	Individual; forced choice	Psychology; social sciences	Quantified subjective data	Generalizability difficult