Decision Sciences
Volume 28 Number 3
Summer 1997
Printed in the U.S.A.

Managing Delphi Surveys Using Nonparametric Statistical Techniques*

Roy C. Schmidt
Department of Information and Systems Management, School of Business and Management, The Hong Kong University of Science and Technology, Clearwater Bay, Sai Kung, Hong Kong, email: schmidt@uxmail.ust.hk

ABSTRACT

Information systems researchers have often turned to a variant of the Delphi survey technique to support their research of key issues in their field. Two particular weaknesses of past studies using this approach have been a lack of a definitive method for conducting the research and a lack of statistical support for the conclusions drawn by the researchers. In this paper, the author presents a method, based on nonparametric statistical techniques, to conduct ranking-type Delphi surveys, perform analysis, and report results. The author takes this one step further by illustrating an actual analysis of a Delphi survey. The analysis is compared to results that were presented without the benefit of the author's approach. This paper shows that use of the advocated approach can streamline and strengthen studies, improve the validity of results, and thus better serve the consumers of the research findings. Since the ranking-type Delphi is so popular among information systems researchers, a consistent method is needed to apply to their data collection, analysis, and reporting of results. This paper provides such a method in concise form and illustrates the use of the method in a manner affording comparison between it and previous practice.

Subject Areas: Group Decision Making, Nonparametric Statistics, Research Design, and Strategic Planning.

INTRODUCTION

A derivative of the Delphi method (Dalkey, 1969; Dalkey & Helmer, 1963) has been used extensively in information systems (IS) research to identify and rank key issues for management action. This Delphi variant has also been the basis for some techniques used in group support systems (GSS) (Valacich, Dennis, & Nunamaker, 1991). In this paper, the goal is to improve the approach that is termed the "ranking-type" Delphi by providing a sound method for the collection and analysis of data.

*The author would like to thank Dan Couger for his kind cooperation in using the data as presented in his previous paper (Couger, 1988b) to illustrate the techniques advocated. Thanks also to a blind reviewer and the associate editor for insightful comments that strengthened and streamlined the paper.

THE DELPHI METHOD

The Delphi method was devised at Rand Corporation in the 1950s (Dalkey & Helmer, 1963) as a means to handle opinions rather than objective facts. The method uses an iterative feedback technique with a group of experts. Forecasting has been a major area of application of the method in many different fields, such as public administration (Preble, 1983), medicine (Spiby, 1988), and technology diffusion (Gray & Nilles, 1983).

In the management disciplines, a modified approach has been used to shape a group consensus about the relative importance of issues (Delbecq, Van de Ven, & Gustafson, 1975). This ranking form of Delphi has been used in diverse fields such as social work education (Ruskin, 1994), operations management (Malhotra, Stelle, & Grover, 1994) and IS (e.g., Brancheau & Wetherbe, 1987).

For forecasting, a set method has been applied across studies (Bordley & Wolff, 1981; Dalkey, 1969; Linstone & Turoff, 1975). But for the ranking-type Delphi, no particular method has been published. Researchers have not followed any consistent method, and have not used uniform means to report results. Beyond the question of how to track consensus, there are three important issues that need to be resolved before the ranking form of Delphi can be considered a sound method.

First, the researcher must know when to stop polling. If the polling stops too soon, the rankings may not be meaningful; but, too many rounds would tax the researcher's resources and waste the panel members' time. Second is how many items to carry over to subsequent rounds. Too many items for ranking can cloud consensus. Third, researchers have not made good use of available statistical techniques to support their conclusions.

In the following sections, a method is presented showing how to pare a list of issues to those few of interest, and how to determine with confidence when the ranking process should be stopped. A straightforward method using nonparametric statistics is illustrated. The method is sound, powerful, and computationally simple. The presentation of results is easy to follow, making them accessible to practitioners as well as academics. This method is discussed in detail and two surveys conducted by Couger (1988a, 1988b) are reanalyzed for comparison of methods.

THE REANALYSIS

Couger's (1988a, 1988b) surveys were chosen because they illustrate the major difficulties in conducting this sort of research. Couger conducted two separate surveys (1988b, p. 162), one concentrating on IS executives and a second using only human resource managers. For the first round of each survey, Couger asked each respondent to list the six most important human resource issues, in rank order, for IS management in the 1990s. The participants described each issue and gave their rationale for putting it on the list. In the second round, respondents were asked to examine a ranked, consolidated list of all issues from the first round, and to submit their ranking of what they considered to be the "top (20) issues" (1988b, p. 162), ties not allowed. Couger transformed these ranks to the top 10, giving the bottom 10 items ranks of 0, before compiling the aggregate, ordered list for the third round. For the third (and final) round, the panelists were then asked to rank their top 10 choices from this list, giving the remaining issues a rank of 0. Again, ties were not allowed.

Couger (1988a, 1988b) compared the mean ranks of the top 10 issues in the second and third rounds for each group to a "perfect consensus" mean ranking by graphic comparison of the mean ranks. He used the size of the "standard deviations" to break ties in mean ranks. He concluded that a consensual top 10 issues had been isolated in each survey. By subjectively matching issues between the two surveys, Couger determined that the two groups had "substantive agreement" (1988b, p. 169) on the key issues.

A summary of the data for Couger's IS directors (1988b) is presented in Table 1. The issue number is the final rank. The number of panelists (k) for each round is given at the top of each column. The columns headed "D^2" show the squared differences between the mean ranks and the grand mean, that is, $D_i^2=(\bar{R}_i-\bar{R})^2$. This is analogous to variance, so it is called "variance of rank."

Couger (1988b) ranked the items from Round 1 by combining percentage of mentions and relative rankings by the individual respondents. As can be seen from the percentage of mentions, none of the items was supported by a majority of the respondents. An approximation of the mean ranks has been produced by multiplying each percentage of mention by its first-round rank as assigned by Couger. One is less interested in the precise order of the items than in the level of agreement. Here, agreement is measured using Kendall's coefficient of concordance (W) (Kendall & Gibbons, 1990; Siegel & Castellan, 1988). Using W, one can make a realistic determination of whether any consensus has been reached, whether the consensus is increasing, and the relative strength of consensus. (In rounds 2 and 3, Couger ranked 20 items as 1 through 10 for the top 10, and 0 for the bottom 10. To compensate for this, a factor was added to the denominator of the formula for W. The modified formula is available from the author.) Some researchers, such as Brancheau and Wetherbe (1987), only used Kendall's W in the final round of their surveys. Table 2 shows a guideline to interpretation of Kendall's W; however, it is not intended to show exact cutoffs.

Kendall's method measures current agreement (the ordered list by mean ranks) with a least squares solution. It is the most popular method for this purpose, mainly due to its simplicity of application. Cook and Seiford (1978) proposed an alternative method based on a least absolute distance solution; however, their method produces multiple solutions that are equally good, making its use in a Delphi survey impractical.

From the Round 1 data, a low level of agreement is found among the judges (see bottom of Table 1). This should be expected given the conditions under which the items were ranked. In Rounds 2 and 3, the raters were moving to higher levels of agreement in each round. Couger (1988b) showed that after the third round, the median ranks for issues 11 through 20 were all 0. From this evidence, it can be said that the IS executives achieved moderate consensus on the top 10 issues. The mean ranking of the third round is a fair estimate of the "true" consensus ranking for the top 10 items (Siegel & Castellan, 1988). Couger (1988b) treated the data from the human resource managers in the same way (Table 3).

The tracking of consensus for the human resource managers is quite different from that of the IS managers. The consensus was already moderate in the first round. Unlike the IS managers, three issues were mentioned by the majority of respondents in the first round, and the top issue was mentioned by 74% of the managers. As a result, the top dozen issues were already apparent in the first round. Looking at

Table 1: Ranking of issues by information systems managers.

	Round 1 (k=57)			Round 2 (k=43)		Round 3 (k=43)	
Issue Number	Percentage Who Mention	Mean Rank	D^2	Mean Rank	D^2	Mean Rank	D^2
1	41.8	7.94	33.60	6.8	16.48	8.60	34.35
2	26.5	4.77	6.89	5.0	5.11	6.60	14.90
3	49.0	9.80	58.60	4.4	2.76	6.07	11.06
4	12.2	2.07	0.01	4.6	3.46	5.58	8.05
5	8.2	0.90	1.54	5.2	6.05	5.58	8.05
6	8.2	0.41	3.01	4.2	2.13	4.56	3.29
7	22.4	3.36	1.48	4.4	2.76	3.86	1.24
8	12.2	1.46	0.46	3.2	0.21	3.37	0.39
9	16.3	2.12	0.00	3.4	0.44	3.05	0.09
10	16.3	2.28	0.02	2.9	0.03	2.37	0.14
11	16.3	2.61	0.21	2.9	0.03	1.60	1.30
12	8.2	0.82	1.76	1.4	1.80	1.05	2.88
13	12.2	0.85	1.67	2.4	0.12	1.02	2.96
14	14.3	1.29	0.74	1.6	1.30	0.81	3.73
15	14.3	1.14	1.00	1.1	2.69	0.35	5.74
16	4.1	0.12	4.09	0.4	5.48	0.14	6.78
17	8.2	0.33	3.30	0.4	5.48	0.09	7.03
18	4.1	0.04	4.43	0.2	6.45	0.09	7.03
19	8.2	0.49	2.73	0.2	6.45	0.07	7.15
20	4.1	0.08	4.26	0.1	6.97	0.0	7.53
Totals		42.90	129.79	55	76.17	55	133.69
Grand Means		2.145		2.75		2.75	
		W	X^2	W	X^2	W	X^2
		.195	211.19*	.33	269.61*	.57	465.69*

*$p < .001$

Couger's data (1988b, Table 2), the group formed a moderate consensus on the ranking of the top 11 issues by giving a median rank of 0 to the bottom nine.

How about the agreement between the two groups? For this question in the present paper, Kendall's rank-order correlation coefficient (T) is used (Kendall & Gibbons, 1990; Siegel & Castellan, 1988). T is used rather than the Spearman rank-order correlation coefficient because it emphasizes the relative ordering of the issues rather than the magnitude of difference between ranks. A one-tailed test of significance is also used. (If the agreement is not significant, or if a significant level of disagreement is found, then a different view will have been established.) Among Delphi researchers, Watson and Brancheau (1991) used Kendall's T to compare subjectively mapped responses from several independently conducted surveys.

Only the final rankings for those items that Couger (1988b) subjectively mapped across both groups (Table 4) were tested. The exact nature of the issues may differ between the two surveys. Eight of these issues were in the top 10 for the

Table 2: Interpretation of Kendall's *W*.

W	Interpretation	Confidence in Ranks
.1	Very weak agreement	None
.3	Weak agreement	Low
.5	Moderate agreement	Fair
.7	Strong agreement	High
.9	Unusually strong agreement	Very high

Table 3: Ranking of issues by human resource managers.

	Round 1 (k=21)			Round 2 (k=19)		Round 3 (k=19)	
Issue Number	Percentage Who Mention	Mean Rank	D^2	Mean Rank	D^2	Mean Rank	D^2
1	74	14.8	115.48	8.3	30.80	8.9	37.76
2	47	8.93	23.78	7.1	18.92	7.5	22.52
3	37	5.55	2.24	5.9	9.92	7.1	18.88
4	26	2.86	1.42	4.0	1.56	4.9	4.60
5	58	9.28	27.31	4.0	1.56	4.6	3.40
6	53	9.54	30.10	4.2	2.10	4.4	2.71
7	47	7.99	15.49	3.8	1.10	4.4	2.71
8	37	5.18	1.27	3.5	0.56	3.2	0.20
9	21	1.89	4.68	3.1	0.12	2.7	0.00
10	11	1.10	8.73	2.1	0.42	1.9	0.73
11	42	5.46	1.98	3.1	0.12	1.5	1.58
12	5	0.15	15.24	1.0	3.06	1.0	3.08
13	42	5.04	0.97	2.7	0.00	0.8	3.82
14	11	0.66	11.52	0.7	4.20	0.7	4.22
15	16	1.28	7.70	0.8	3.80	0.4	5.55
16	5	0.25	14.47	0.0	7.56	0.4	5.55
17	11	0.77	10.78	0.6	4.62	0.3	6.03
18	5	0.20	14.85	0.1	7.02	0.2	6.53
19	5	0.10	15.63	0.0	7.56	0.1	7.04
20	5	0.05	16.03	0.0	7.56	0.1	7.04
Totals		81.08	339.67	55	112.61	55	143.95
Grand Means		4.054		2.75		2.75	
		W	X^2	W	X^2	W	X^2
		.51	203.80*	.48	173.28*	.62	223.82*

*$p < .001$

Table 4: Comparison of ranks between groups.

Issues	IS Manager's Rank	HR Manager's Rank
Business orientation	1(&5)	3
End-user focus	2	11
Retraining	3	6
HR planning	4	1
Creativity	6	14
Interpersonal skills	7	13
Motivation	8	4
Rewards	10	5
Dual career ladder	11	15
Value shifts	12	8

IS executives, while 6 were in the top 10 for the human resource executives. The table is arranged in the rank order assigned by the IS managers. (Either group's ranking can be used without affecting the statistic.)

Here, it is found that T=.244. Consulting a table of exact probabilities for T (Seigel & Castellan, 1988, p. 362), the one-tailed probability is p=.19. So, the two groups of managers do not agree on the ranks of these items. Couger also argued that there was "considerable agreement" in placement of five key issues in the top 10 by both groups (1988b, p. 169). This assertion can be tested using the Fisher Exact Test, which is tabled (Seigel & Castellan, pp. 335-338). The arrangement in Table 5 is not significant (p=.667). Therefore, there is no relation between the two groups' categorization of the issues.

The necessary statistics for the survey data have been computed and some preliminary observations on the implications of these statistics have been made. Now the mechanics of a new method and the use of these statistics to illustrate the utility of the method are presented.

OUTLINE OF METHOD

In this section, a new method for ranking-type Delphi surveys is outlined, covering three phases. The first topic is how to conduct and manage data collection (the survey) to make the best use of resources. This is followed by an analysis of the data and a discussion of presentation of the data to support research findings.

Managing Data Collection

There are three distinct phases in data collection: (1) the discovery of issues, (2) determining the most important issues, and (3) ranking the issues.

Phase 1

To maximize the chance of unearthing the most important issues, the respondents should be encouraged to submit as many issues as possible in this first phase. Researchers have typically limited the number of items a single respondent can contribute (e.g., Couger, 1988b). It is better to ask the respondents to list at least

Table 5: Grouping of issues.

	IS Managers' Top 10		
HR Managers' Top 10	Yes	No	Totals
Yes	5	1	6
No	3	1	4
Totals	8	2	10

six important issues. Having the respondents describe each issue is essential at this stage, because several respondents are likely to raise the same issue using different terms.

The researcher then consolidates the responses into a single list. Where several different terms are used for what appears to be the same issue, the researcher should list all the terms together, then provide one consolidated description of the issue. Next, the respondents verify that the terms have been properly mapped and that their ideas have been fairly represented. Without this step, there is no basis to claim that a valid, consolidated list has been produced. Previous researchers have not taken this precaution. If the research plan includes comparing the responses of one or more groups, then all groups must participate collectively in the first phase. Otherwise the researcher will face great difficulty in subjectively mapping the groups' independently ranked items for comparison.

If major differences are uncovered by the respondents, the researcher may need to repeat this step. Once the differences are resolved, the survey enters the second phase.

Phase 2

Now the researcher must pare the list of issues so that they can be meaningfully ranked. Researchers have typically combined this phase with the third phase. But the sheer number of issues involved hinders the ranking exercise. As each respondent independently ranks some arbitrary number of top issues, the mean ranks usually show little variation. Furthermore, the arbitrary choice of the number of issues to be ranked is not a good way to bound the final list.

As a dramatic example, consider Couger's second survey (1988b). In the second round, he asked respondents to rank the top 20 issues from a consolidated list. But the respondents had already reached consensus on the top 12 to 15 issues in the first round. As a result, the level of consensus actually declined slightly, and the issue finally ranked as number 13 was temporarily in the top 10. The list should be bounded statistically rather than arbitrarily.

A randomly ordered, consolidated list from the first phase is sent to each participant. They independently select at least 10% (or more if the list contains less than 100 items) of the issues as the most important. The researcher then eliminates all issues that were not selected by a simple majority of the respondents.

In past surveys, respondents usually agreed on a list of less than 20 items rather quickly. So, it is suggested that if the pared list consists of around 20 items, the researcher should go on to the third phase. If the pared list is much larger than 20

items, a second round of Phase 2 could be conducted, using the shortened list. If this second round does not result in a reasonable reduction in the number of items, the researcher must make a tough decision to either abandon the survey or to arbitrarily pare the list. It is not the researcher's place to decide the top issues, but such arbitrary reduction in list size has been the norm.

The resulting list may contain any number of issues. Setting an arbitrary size for the list forces the result. In Table 3, the issues with a second-round mean rank of 0 gained in mean rank in the third round. This suggests that these issues distracted the panel from concentrating on the "top 10." Conversely, there may be additional issues that the respondents consider important that are not included because they do not make the "top 10."

If the study compares the responses of two or more groups (as in Brancheau & Wetherbe, 1987; Couger, 1988b; Watson & Brancheau, 1991), the groups should be separated beginning with the second phase. The groups develop a common list of issues with a common set of definitions in the first phase. Then in the second phase, they can diverge in composing the list for ranking. Couger did not begin with a common list, so he was forced to map the issues subjectively. Watson and Brancheau avoided this problem in their comparison of four independently conducted surveys by repeated use of a list of issues generated in the first survey. Brancheau and Wetherbe contrasted their two groups by a post hoc separation of responses. In the latter two studies, undesired influence in choice of issues and in final rankings, respectively, weakened the outcomes.

Phase 3

Next, the pared list is arranged in random order, and the respondents are asked to rank all the issues. The researcher can compensate for tied rankings when computing Kendall's W, but the computation is simpler and the results clearer if the respondents are asked to avoid ties.

Combining individual rankings to produce a consensual ranking has been studied extensively since the 18th century. A survey of methods can be found in Hwang and Lin (1987). Arrow (1951) showed that no method can produce a consensual choice not influenced by the method. Kendall's method, if used in a one-shot ranking exercise, is no exception. The solution (mean ranks) is imposed by the method and does not necessarily reflect the best solution from a social choice perspective. But in Delphi, the iterative approach lessens this effect by allowing panelists to revise their choices. Kendall's method is then preferable to other methods because it provides a unique solution that is easy to understand and simple to apply.

A special feature of the Delphi approach is controlled feedback to the respondents each round (Dalkey, 1969). Dalkey stressed that more than just the mean should be conveyed to the experts. Past practice in ranking-type Delphis has not followed this advice.

Besides the mean rank, there are three relevant pieces of feedback that can be provided. The first piece is an interpretation of Kendall's W from the previous round. For example, if W is not significant, then the respondents can be told, "There was no agreement among the panel members on the relative ranking of these items in the previous round." The second piece is the percentage of respondents placing each

item in the top half of their list. These two pieces of information convey a sense of the degree of consensus. The third piece is relevant comments by the panelists.

Some researchers have argued that the "standard deviation" of ranks can indicate consensus (e.g., Couger, 1988b; Brancheau & Wetherbe, 1987). The concept of standard deviation does not apply to ordinal level data. There are no fixed intervals between ranks and no absolute reference point to calibrate ranks between panelists. Providing such data to the experts, or using it in research reports, is misleading.

After each round in this phase, the researcher must ask, "Should another round be conducted to obtain greater consensus?" The trade-off between feasibility (the indulgence of the panelists, the researcher's resources, and the additional time required) and the potential gain to be achieved must be considered (see also Nelms & Porter, 1985). When Kendall's W is small, the decision is trivial. But with moderate consensus, the trade-off is more difficult to weigh.

There are two statistical criteria for potential gain that can be used as stopping rules. The first criterion is strong consensus. In the absence of strong consensus, a leveling off of W would indicate a lack of progress from the previous round, so the polling should stop. Further rounds would allow the "pull of the mean" to exert undue influence, giving spurious results.

The statistical significance of W is not sufficient criterion to halt the survey. For panels of more than 10 experts, even very small values of W can be significant. One must only consider the actual value of W as indication of the strength of the consensus.

Previous surveys have been halted when consensus was at a moderate level (e.g., Brancheau & Wetherbe, 1987; Couger, 1988a, 1988b). Since steady progress was being made in these surveys, the decision to halt must have been due to feasibility concerns. Barring such issues, the surveys should have continued for one more round.

Analysis of Results

The procedures for analysis of results, including the calculation of the appropriate statistics and the use of those statistics to manage the conduct of the survey have already been demonstrated. Also, past practice has been critiqued by using these statistics.

For both of Couger's studies (1988a, 1988b), it is believed that one more round of polling would have been appropriate. The order of ranking for many of the issues shifted from round to round, but Couger assumed that in moving to a "perfect" consensus the ranks would not shift again. A fourth round might have produced further changes and could have strengthened the level of consensus.

It was also found that Couger's "substantive agreement" (1988b, p. 169) was illusory. There was no agreement between the two groups on which issues belonged in the "top 10," nor on the ranking of the issues. So the second poll of the human resource managers did provide significant new information on the key issues, contrary to Couger's findings.

Presentation of Results

Numerous approaches have been used to present the results of ranking-type Delphi surveys. A graphical approach has been used (e.g., Couger, 1988b; Malhotra et al.,

1994) to make the results easily comprehensible for the practitioner. Others have used tables presenting whole ranks, mean ranks, "standard deviation," percentage of support per item, and other descriptive statistics. In this reanalysis, the final whole ranks, mean ranks for each round, and what was termed "variance of ranks," along with Kendall's W are shown. Of course, if the readership is interested in relative ranks, then the most important of these are the final whole ranks.

Researchers should report the total number of issues generated in Phase 1, and the strength of support for the issues that were selected as the pared list at the end of Phase 2. The panelists have at least agreed that out of a relatively large list of issues, there are a few that are most important. Then, the readers must understand the level of confidence they can place in the final rankings. For example, a moderate level of consensus does not warrant high confidence in the relative standing of the issues. Also, one must not affix any meaning to the distance between mean ranks. Some panelists may believe the first two issues to be much more important than all the others, or that the last six issues are about the same level of importance.

Next, researchers should report round-by-round levels of consensus and any other measures of association pertinent to the research hypotheses. Sufficient raw data should be provided in the article or appendixes to support the calculation of the statistics.

Finally, the panel of experts must be well described. The researcher should provide demographics such as those shown by Brancheau and Wetherbe (1987) so that the reader can judge the relevance and reliability of the respondents. The response rate for the initial call for participants should be reported, since a low response rate might indicate that potential participants do not consider the ranking exercise relevant or important. Also, the number of panelists for each round must be reported so that (1) the reader can confirm relevant statistics and (2) indications of flagging interest in the survey are brought out. All of this information helps the reader determine the degree of confidence to place in the findings.

CONCLUSION

In the preceding sections, a method for the conduct of Delphi ranking-type surveys, the analysis of data collected in the surveys, and the presentation of the data to support research findings was outlined. It was shown that previous research could have been improved by the application of the method proposed here.

At the beginning of the paper three critical issues were raised. These three issues were answered as follows:

1. A rule for stopping the poll was illustrated using group consensus.
2. An objective method for paring down a list for ranking was proposed.
3. Appropriate statistics were applied to each step of the collection and analysis processes to strengthen the method.

It is believed that the method presented in this paper answers some of the criticisms of past Delphi-like research. It provides rigor where arbitrary practices were the norm. It offers consistency of application and uniform presentation of results, which contribute to the accumulation of knowledge across studies. Researchers are

encouraged to make use of this method and to help improve the method. [Received: November 16, 1994. Accepted: December 10, 1996.]

REFERENCES

Arrow, K. J. (1951). *Social choice and individual values*. New York: Wiley.

Bordley, R. F., & Wolff, R. W. (1981). On the aggregation of individual probability estimates. *Management Science, 27*(8), 959-964.

Brancheau, J. C., & Wetherbe, J. C. (1987). Key issues in information systems management. *MIS Quarterly, 11*(1), 23-45.

Cook, W. D., & Seiford, L. M. (1978). Priority ranking and consensus formation. *Management Science, 24*(16), 1721-1732.

Couger, J. D. (1988a) Key human resource issues in IS in the 1990s. *Proceedings of the Hawaii International Conference on Systems Sciences*, 162-170.

Couger, J. D. (1988b) Key human resources issues in IS in the 1990s: Views of IS executives versus human resource executives. *Information and Management, 14*(4), 161-174.

Dalkey, N. C. (1969). *The Delphi method: An experimental study of group opinion*. RM-5888-PR. Santa Monica, CA: Rand Corporation.

Dalkey, N. C., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science, 9*, 458-467.

Delbecq, A. L., Van de Ven, A. H., & Gustafson, D. H. (1975). *Group techniques for program planning: A guide to nominal group and Delphi processes*. Glenview, IL: Scott-Foresman.

Gray, P., & Nilles, J. M. (1983). Evaluating a Delphi forecast on personal computers. *IEEE Transactions on Systems, Man, and Cybernetics, 13*(2), 222-224.

Hwang, C.-L., & Lin, M-J. (1987). *Group decision making under multiple criteria: Methods and applications*. Berlin: Springer-Verlag.

Kendall, M. G., & Gibbons, J. D. (1990). *Rank correlation methods*. London: Edward Arnold.

Linstone, H. A., & Turoff, M. (1975). *The Delphi method: Techniques and applications*. Reading, MA: Addison-Wesley.

Malhotra, M. K., Stelle, D. C., & Grover, V. (1994). Important strategic and tactical manufacturing issues in the 1990s. *Decision Sciences, 25*(2), 189-214.

Nelms, K. R., & Porter, A. L. (1985). EFTE: An interactive Delphi method. *Technological Forecasting and Social Change, 8*(1), 43-61.

Preble, J. F. (1983). Public sector use of the Delphi technique. *Technological Forecasting and Social Change, 23*(1), 75-88.

Ruskin, M. S. (1994). The Delphi study in field instruction revisited: Expert consensus on issues and research priorities. *Journal of Social Work Education, 30*(1), 75-89.

Siegel, S., & Castellan, N. J., Jr. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.

Spiby, J. (1988). Advances in medical technology over the next 20 years. *Community Medicine, 10*(4), 273-278.

Valacich, J. S., Dennis, A. R., & Nunamaker, J. F. (1991). Electronic meeting support—The GroupSystems concept. *International Journal of Man-Machine Studies, 34*(2), 261-282.

Watson, R. T., & Brancheau, J. C. (1991). Key issues in information systems management: An international perspective. *Information & Management, 20*(3), 213-223.

Roy Schmidt is an assistant professor of information systems in the School of Business and Management, The University of Science and Technology, Hong Kong. He has a B.A. and B.S. from the University of Maryland, an M.B.A. from St. Mary's University, San Antonio, and a Ph.D. in MIS from Indiana University. Before entering academia, Dr. Schmidt served in the U.S. Air Force for over 22 years. His research interests include support of the strategic decision process, executive information use, information systems project management, and information systems project risk management.