APPLYING DELPHI METHODOLOGY IN FAMILY THERAPY RESEARCH

David A. Jenkins
Thomas Edward Smith

ABSTRACT: The Delphi technique, which is increasingly seen in family therapy publications, encourages the articulation of refinement in family therapy practice and theory. Combining both quantitative and qualitative methodology, this technique involves consideration of information that might otherwise have been overlooked. Due to lack of clarity on this approach, a dilemma exists in applying this research procedure in family therapy. To remedy this deterrent, this paper reviewed 11 studies and outlined use of the technique in regard to design, procedure, and sampling.

KEY WORDS: Delphi; family therapy research; methodology.

The Delphi technique, which is increasingly seen in family therapy publications (Avis, 1986; Figley & Nelson, 1989, 1990; Jenkins, 1992; Nelson & Figley, 1990; Nelson, Heilbrun, & Figley, 1993; Stone Fish, & Piercy, 1987; Tester, 1992; Watson, 1985; Wheeler, 1985; Winkle, Piercy, & Hovestadt, 1981), allows for a systematic approach in uncovering information. This approach provides the opportunity to investigate data that might be overlooked with other methodologies. However, Delphi procedures are omitted as an approach in general research text discussions and, as a consequence, these procedures are difficult to apply in social science research in general and family ther-

David A. Jenkins, PhD, is an assistant professor in social work at Texas Christian University, PO Box 30790, Fort Worth, TX 76129. Thomas Edward Smith, PhD, is an associate professor in social work at Florida State University, School of Social Work, R-91, Tallahassee, FL 32306. Requests for reprints should be sent to the first author.

Contemporary Family Therapy, 16(5), October 1994

apy research specifically. To remedy this dilemma, this paper will review and outline use of the Delphi technique as a research methodology in family therapy.

THE DELPHI TECHNIQUE

Researchers using a Delphi technique elicit opinions of knowledgeable individuals who can help to resolve some area of concern. After identifying and defining his or her area of research, the investigator must determine what type of expertise would be required to resolve the problem. Experts are selected for their familiarity with the topic being studied and the information they bring to the project (Riggs, 1983). The researcher then prepares a questionnaire that is distributed to each expert in an initial "round" or "wave" of inquiry (Avis, 1986; Jenkins, 1992; Stone Fish, 1985). The questionnaire states the area of concern and requests the expert's opinions. Next, the researcher analyzes the material received from the first questionnaire to determine whether the dilemma has been resolved through the first round of investigation.

If there has not been enough refinement of the area being studied, the researcher mails the same questions back to participants with their responses from the first round of inquiry (Preble, 1983; Riggs, 1983). This process of mailing questionnaires, analyzing results, and reformulating the questionnaire with the previous round comments is continued until opinion is further refined or consensus is reached on the subject area.

Helmer and Rescher (1959) established a theoretical background for the Delphi technique when they suggested that resolving dilemmas called for the judgment of several experts. In the absence of a convincing reason to select a particular plan of action, experts could be used to help end indecision (Linstone & Turoff, 1975). The Delphi technique offers an effective use of group information (Dalkey, 1972) by allowing participation without problems of social inhibition (Dalkey, Rourke, Lewis, & Snyder, 1972). By gathering and refining expert opinion, the technique makes use of speculation to forecast trends, build consensus, and assess current need in regard to or concerning a given area of concern (Dalkey, 1972; Redenour, 1982).

The Delphi technique was originally used to help predict horse race outcomes (Preble, 1983). However, the RAND Corporation adapted the technique for assistance in making defense and military decisions in the 1940s (Dalkey & Helmer, 1963). It continued to be

used throughout the 1960s to address military concerns, and later began being used to forecast both professional and business events (Preble, 1983). Additionally, the Delphi has proven effective in forecasting future events (Riggs, 1983). Dalkey (1969), asserts:

> I can state from my own experience, and also from the experience of many other practitioners, that the results of a Delphi exercise are subject to greater acceptance on the part of the group than are the consensuses arrived at by more direct forms of interaction (p. 17).

Not only has the Delphi technique been used in public sector settings for purposes of forecasting, public budgeting, and goal setting (Preble, 1983), but Helmer (1967) notes that it has an important use in model building. In other words, it is used when there is an absence of an accepted theoretical body of knowledge to point to a particular decision (Helmer, 1967). For example, Figley and Nelson (1989) used the Delphi technique to identify the most important characteristics of a beginning marriage and family therapist. In this approach, as compared to other types of polling procedures, the focus is more on identifying issues and "exploring minds" than with setting down precise recommendations (Wheeler, 1985).

Advantages of the Delphi Technique

The Delphi technique offers several advantages that are not apparent or available in other types of analogous consensus building processes (Figley & Nelson, 1988; Sackman, 1975). One advantage is that it allows respondents to remain anonymous (Brown, Cochran, & Dalkey, 1969). This anonymity enables respondents to react without coercive efforts from other participants. Instead of attempting to refine ideas through face-to-face discussion, committee meetings, and confrontation, the Delphi process allows for individual input and controlled opinion feedback (Helmer & Rescher, 1959).

Additionally, because of the great geographical distance that exists between many of the participants, discourse that might otherwise have been logistically impossible occurs in this method of study (Figley & Nelson, 1990).

With the Delphi technique, each participant is given time to consider his or her response in ways that might not have been possible in group decision-making meetings.

Additionally, this technique allows the researcher to manage re-

turns from participants, while providing feedback in the form of responses and analysis (Brown, Cochran, & Dalkey, 1972). This feedback allows participants to consider other people's responses and determine if those responses coincide with their own opinion on the topic.

In a developing field like family therapy, Delphi studies allow efficient and rapid collection of expert opinions. The anonymity of a Delphi process diminishes social desirability response sets and encourages refinement of opinion on critical topics. It also provides an effectual method for conducting preliminary research prior to beginning a large-scale study (Dalkey & Helmer, 1963).

Disadvantages of the Delphi Technique

However, Delphi results are only as good as their methodology. As Dalkey (1969) states:

> Like any technique for group interaction, the Delphi procedures are open to various misuses; much depends on the standards of the individual or group conducting the exercises (p. 17).

Without careful scrutiny of methodological procedures, Delphi studies can be reduced to nonprobabilistic samples with unvalidated measures that receive little careful analysis. An important consideration in working with the Delphi technique is that this method contains elements of both quantitative and qualitative research techniques. This combination has made it difficult for researchers to formalize inquiry procedures. The RAND Corporation has been instrumental in the development of several books and articles that outline how Delphi procedure was created and refined (Brown, Cochran, & Dalkey, 1969; Dalkey, 1969; Dalkey & Helmer, 1963; Dalkey, Rourke, Lewis, & Snyder, 1972). However, these do not offer the specific detail needed for applying this technique in family therapy. With this in mind, the remainder of this paper outlines use of the Delphi procedure as found in family therapy studies.

SUMMARY OF FAMILY THERAPY DELPHI STUDIES

The following review examines key aspects found in the methodologies of 11 Delphi studies conducted in family therapy (Avis, 1986; Figley & Nelson, 1989, 1990; Nelson & Figley, 1990; Nelson,

Heilbrun, & Figley, 1993; Jenkins, 1992; Stone Fish, 1985; Tester, 1992; Watson, 1985; Wheeler, 1985; Winkle et al., 1981). It should be noted that each Delphi study was examined only as completely as the manuscripts and dissertations allowed. Authors' words were used as much as possible to capture the researchers' intent; however, some discussions omit critical detail in the description of their study. The Delphi methodology will be evaluated in regard to the following issues: 1) design, 2) procedure, and 3) sampling.

DESIGN

Careful design of any Delphi study incorporates several fundamental aspects, including: 1) the number of rounds or waves used, and 2) questionnaire construction.

Number of Rounds or Waves

Most Delphi studies in the field of family therapy appear to be designed with either two or three rounds or waves (i.e., the times the panelists respond to the questionnaire and items in the study). Table 1 provides a detailed summary of the number of rounds used in each of the 11 family therapy Delphi studies.

Although Delphi designs with two rounds are methodologically acceptable (Martino, 1972), some family therapy authors list specific justifications for two-round studies (Avis, 1986; Jenkins, 1992; Wheeler 1985). The studies by Wheeler (1985) and Avis (1986) were originally slated to have three rounds, but ended after two rounds. Even though Jenkins (1992) included a rationale as to why his study might continue beyond two rounds, it stopped at two. These authors state they discontinued the inquiry because adequate consensus had been reached, and there was the possibility of participants' fatigue which might hinder the project. Unfortunately, the reader is left unclear how authors determined that adequate refinement or consensus had been reached and that the next round of inquiry was not needed. These studies implied that the next round was not conducted because little new information was being added to the study and consensus had been reached on at least one item (Avis, 1986; Jenkins, 1992; Stone Fish, 1985). However, standards set in qualitative research methods (e.g. Goetz & LeCompte, 1984; Miles & Huberman, 1984) offer researchers a guide to determine when "refinement" is beginning and there is no need for further rounds of inquiry.

TABLE 1
Number of Rounds in Family Therapy Delphi Studies

Winkle and associates (1981)	3
Stone Fish (1985)	3
Watson (1985)	3*
Wheeler (1985)	2
Avis (1986)	2
Figley & Nelson (1989)	2
Figley & Nelson (1990)	2
Nelson & Figley (1990)	2
Nelson, Heilbrun, & Figley (1993)	2
Jenkins (1992)	2
Tester (1992)	2*

*An additional round was used to solicit comments from the panelists.

The follow-up studies in the Basic Family Therapy Skills Project (Figley & Nelson 1990; Nelson & Figley 1990; Nelson, Heilbrun, & Figley, 1993) built upon an original two-round Delphi study by Figley and Nelson (1989). Those studies were used to refine a specific area of the data obtained in the original project by adding an additional round of inquiry. The use of follow-up studies could be used in other Delphi studies to further refine data obtained in two-round studies.

A slight shift in the traditional two-round study was done by Watson (1985) and Tester (1992). They incorporated pre-existing definitions of their researched topic as part of their questionnaire in the first round of inquiry. They argued that this process enabled them to start the study with a second round inquiry since participants were already responding to information upon receiving the first questionnaire. Therefore, a study that used only two rounds of inquiry was said, in effect, to have three.

It is unclear how use of pre-existing definitions affects a Delphi project; however, it is possible that these definitions enabled participants to respond with less possibility of fatigue. Riggs (1983) claims that using preexisting information helps reduce the number of rounds in a study, thus saving time. One negative consideration is that this information may bias or limit how the experts respond to the items on the questionnaire. However, one has to assume that "knowledgeable participants" will be familiar with the literature and will be affected by preexisting definitions whether they appear on the first question-

naire or not. As an example, Jenkins (1992) stated that many of his first round participants asserted that the answer to the original questionnaire questions were already in the literature and merely cited articles as their response.

Questionnaire Construction

Most Delphi studies begin with several open-ended questions that elicit information from respondents. The first questionnaire can contain any number of questions or statements, but the majority of family therapy studies have used three to nine questions in the initial round of inquiry (Avis, 1986; Jenkins, 1992; Stone Fish, 1985; Tester, 1992; Watson, 1983; Wheeler, 1985; Winkle et al., 1981). Some researchers gave panelists a "blank page" with a single question or statement in the first round (Figley & Nelson, 1989), or used the final results in a previous study to outline their questionnaire (Figley & Nelson, 1990; Nelson & Figley, 1990).

As stated earlier, in addition to open-ended questions, Tester (1992) and Watson (1985) included pre-existing definitions on the first questionnaire. Including definitions on the first questionnaire made part of the questionnaire have existing answers, and thus the questionnaire elicited both feedback and original insight from experts. Although most researchers use fairly unstructured categories of questions on the initial questionnaire, these questions limit what is and is not researched, and, therefore, affect study results. Table 2 shows the number of questions included on the first round of questionnaires for each of the family therapy Delphi studies.

Since the questions or prompts used in the first wave are critically important, a pilot study may be used to help construct the remainder of the study. The pilot shapes the eventual selection and development of the categories used in the questionnaire (Jenkins, 1992; Riggs, 1983; Tester, 1992). The second (and later) round questionnaire is devised from the material collected from the previous round of inquiry. Most researchers strive to condense the data obtained in the first round into non-redundant and non-overlapping items.

Additionally, a 5- or 7-point Likert-type scale is placed beside each item for participants to indicate their level of support. The Likert scale is anchored on one end with a term such as "strongly agree," while the other end is anchored with "strongly disagree." In this format, participants rate their degree of acceptance or disagreement with each item on the questionnaire. This questionnaire wave is usually mailed only to panelists who participated in the first round of study.

TABLE 2
Number of Categories or Questions on the First Questionnaire

Winkle, Piercy, & Hovestadt (1981)	9*
Stone Fish (1985)	8
Watson (1985)	3
Wheeler (1985)	7
Avis (1986)	8
Figley & Nelson (1989)	2
Figley & Nelson (1990)	**
Nelson & Figley (1990)	**
Nelson, Heilbrun, & Figley (1993)	**
Jenkins (1992)	6
Tester (1992)	9

*Nine categories were listed in tables in their published report; therefore, one has to assume that there were approximately nine open-ended categories or questions on the first questionnaire.
**These studies were continuations of the first study in 1989; therefore, these studies did not actually have a first round questionnaire that was separate from the first study.

If a third round is deemed necessary, the questionnaire is constructed of the summary score (i.e., median or mean) that was reached on each item and a Likert-type scale. Once again, with this additional information, participants are asked to indicate whether they endorse or reject the items on the survey. If further rounds are deemed necessary, the questionnaire is once again sent with the collected statements and summary scores from the previous round. The process of additional waves may continue until the researcher determines adequate refinement has been reached or until it is no longer advantageous to continue.

To cope with the volume of responses that may be received, some authors divided the material received from the initial round of study into several smaller, more manageable questionnaires (Figley & Nelson, 1989; 1990; Nelson & Figley 1990; Nelson, Heilbrun, & Figley, 1993). These smaller questionnaires were then randomly sent to panelists. Figley and colleagues (1989) adopted this strategy to reduce panelists' fatigue by lessening the magnitude of effort from any one group. Although a creative deviation from other Delphi studies, this multi-stage sampling is deemed acceptable for qualitative inquiry (Figley & Nelson, 1989).

PROCEDURE

The importance of the coding system used to formulate the second questionnaire from the first round data and the analysis used to interpret the results from the second and following rounds cannot be overlooked in Delphi methodology.

Coding System

Delphi researchers have had the difficult task of developing a coding system to synthesize the massive amounts of data obtained in the first round of their study into a manageable second round questionnaire. The criteria the authors choose to formulate responses in the second questionnaire is a critical issue. In an attempt to use the panelists' wording as closely as possible, the coding system used in many of the studies consisted of the researcher summarizing, editing, and eliminating duplicate or redundant items (Avis, 1986; Figley & Nelson, 1989, 1990; Jenkins, 1992; Nelson & Figley, 1990; Nelson, Heilbrun, & Figley, 1993; Stone Fish, 1985; Wheeler, 1985). This coding system allowed the researcher to put the summarized material into categories for the next round of the study.

Additionally, several authors used independent raters as part of their coding procedure (Avis, 1986; Jenkins, 1992; Stone Fish, 1985, Tester, 1992; Wheeler, 1985). Those authors used reliability checkers to ensure that concepts included in the first questionnaire were accurately represented in the subsequent questionnaire. Independent raters were chosen for their knowledge of the material or the objectivity they brought to the study. Several authors included a reliability score to indicate the percentage of items that were correctly included in constructing the second questionnaire when independent raters checked the data (Jenkins, 1992; Stone Fish, 1985; Tester, 1992).

Other researchers used their own judgement to transform the first round information into the second questionnaire (Figley & Nelson, 1989, 1990; Nelson & Figley, 1990). This "inductive method" of coding was used to reduce the data into some manageable amount for the panel to consider (Figley & Nelson, 1989). In these studies, one author grouped the material into categories, while the second author made judgments about similarities and groupings. The first author checked the groupings and made minor suggestions and changes until a consensus was reached by both authors. The authors stated

that they attempted to overcome bias by using the wording of panelists rather than attempting to interpret, extrapolate, or infer meaning in coding for the next questionnaire. Additionally, these authors said that this inductive method was consistent with qualitative research methods (Figley & Nelson, 1989).

Two researchers did not manually decide how the data would be condensed into the next questionnaire. Rather, they conducted a content analysis of the data to form the responses for the second round of the survey (Tester, 1992; Watson, 1985). This analysis allowed the researcher to form composite definitions from the data and use them as the basis of the next questionnaire. In this approach, the data are able to "speak for itself," rather than forcing the formation of categories that may or may not fit with the information obtained (Tester, 1992).

Analysis

After gathering, sorting, and condensing information a Delphi researcher analyzes the obtained data. Analysis is used to enable the researcher to accomplish two tasks. First, it allows him or her to obtain a summary score for each item in the study. Second, it allows the researcher to have a means to form comparisons and draw conclusions in order to construct the final profile of the topic. Table 3 offers a synopsis of the types of analytic methods used in family therapy Delphi research.

The most common forms of analysis in family therapy Delphi research are the use of medians and interquartile ranges (Avis, 1986; Jenkins, 1992; Stone Fish, 1985; Wheeler, 1985; Winkle et al., 1981). In these studies, panelists endorsed items for the final profile of the study using a 5- or 7-point Likert-type scale. After retrieving the questionnaires, the authors calculated a median and interquartile range for each item being considered by panelists. Scores of the median and interquartile range help to eliminate opinions that deviate greatly from the majority (Riggs, 1983). When there were more than two rounds in the study, panelists were shown medians and interquartile ranges on the questionnaire when they were asked to respond to the items.

Presumably, the descriptive statistics offered panelists additional information in deciding about an item's acceptability. For the researcher, a predetermined range for acceptable and unacceptable medians and interquartile ranges was calculated to determine whether

TABLE 3
Analytic Strategies

	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*
Median	X	X		X	X				X	X	
Interquartile Range	X	X		X	X				X		
Mean						X	X	X			X
Standard Deviation						X	X	X			X
Themes/Content Analysis			X							X	
Voting			X								
Interviews										X	

Key for Table:
A = Winkle, Piercy, & Hovestadt (1981)
B = Stone Fish (1985)
C = Watson (1985)
D = Wheeler (1985)
E = Avis (1986)
F = Figley & Nelson (1989)
G = Figley & Nelson (1990)
H = Nelson & Figley (1990)
I = Jenkins (1992)
J = Tester (1992)
K = Nelson, Heilbrun, & Figley (1993)
NA = not available or not applicable

items would be included in the final profile of the study. Usually, based on guidelines from previous studies (e.g. Binning, Cochran, & Donatelli, 1972), it is unclear how this acceptable range was first determined. A common guideline score for items to be considered "acceptable" for those studies was a median of less than or equal to 2 (or between 6 and 7, depending on the way the numbers are anchored) and an interquartile range score of less than or equal to 1.5 (Avis, 1986; Jenkins, 1992; Stone Fish, 1985; Wheeler, 1985; Winkle et al., 1981). A median score of 1 or 2 meant that the item was deemed "most important" or that panelists had "strong agreement" with the item being included in the study. Additionally, an interquartile range of less than 1.5 meant that a majority of the scores did not greatly differ from one another.

Moving away from the traditional combination of medians and interquartile ranges, Tester (1992) chose to combine aspects of both qualitative and quantitative analysis. Although still using medians, she eliminated the use of interquartile ranges because it grouped data into arbitrary intervals. Instead, she employed content analysis and interviews for additional analysis. The content analysis was used to group the data, while interviews were done at the end of the study to elicit unstructured comments about the research project. Similarly,

Watson (1985) contacted panelists to have them vote on a favorite definition as the project neared completion. The use of these qualitative types of analysis allowed researchers to contemplate information that might have been overlooked with interquartile ranges (Tester, 1992).

Four Delphi studies used a mean and standard deviation to analyze the items in their study (Figley & Nelson, 1989, 1990; Nelson & Figley, 1990; Nelson, Heilbrun, & Figley, 1993). These authors calculated a mean and standard deviation for each item that was on the second round questionnaire. At the conclusion of the study, the items with the lowest means were presented, along with their standard deviation, in the final profile.

An additional field of analysis delineated in several family therapy Delphi studies is the area of greatest conflict or disagreement. Authors stated they wanted to include this area of conflict or disagreement since their topic of research was in a formative stage, and this additional data offered valuable information (Avis, 1985; Jenkins, 1992; Wheeler, 1985). Additionally, this information is considered significant because authors (Linstone & Turoff, 1975; Sackman, 1975) suggest that items that typically are part of a final profile represent compromise and lack the significance that extreme positions might possess. Therefore, the items considered controversial would offer additional information that could be considered in later studies.

For clarification, researchers need to be clear where the area of "disagreement" is formed on most Delphi Likert-type scales. While scores at one end of the Likert-type scale form the area of "acceptance" for the final profile, the scores at the opposite end of the Likert-type scale form the area determined by agreement not to be included in the final profile. The median scores that are congregated at the center of the Likert-type scale form the area of "disagreement" or "controversy".

Faced with many options for analytic procedures, authors appear to need a reason to select one strategy over another. Even though a mean or median can be used to show central tendencies, medians are considered most useful in describing data that may be sensitive to outliers and might otherwise be distorted with a mean (Cohen & Holliday, 1982). Most authors who use a median as a measure of central tendency also use an interquartile range, or quartile deviation, as a measure of variability. This score shows the "spread" of the scores around the median.

The interquartile range reduces the influence of extreme scores and shows patterns of quartile; one quartile deviation on either side of the median in a normal distribution contains 50% of the cases (Cohen & Holliday, 1982). Tester (1992) disagreed about the utility of this measure of variability and stated that an interquartile range arbitrarily groups data. She proposed that eliminating interquartile ranges still allowed for examination of the data without any loss of meaning. Tester dropped the interquartile range scores that typically had been used to show the "spread" of the scores around the median and included other analysis to help describe her data.

SAMPLING

In order to investigate any field with a Delphi methodology, "elite" or knowledgeable interviewees must be obtained. Operationalizing criteria for these "elite" panelists is critically important, as is locating and utilizing them for the study. In the family therapy Delphi studies, authors chose their sample on the basis of such characteristics as publications, presentations, and/or years teaching or practicing clinical work on the topic area (Avis, 1986; Jenkins, 1992; Stone Fish, 1985; Tester, 1992; Watson, 1985; Wheeler, 1985), administrative positions in graduate-level marital and family therapy training programs (Winkle et al., 1981), and membership in the American Association for Marriage and Family Therapy or the American Family Therapy Academy (Figley & Nelson, 1989; Figley & Nelson, 1990; Nelson & Figley, 1990; Tester, 1992). Table 4 summarizes the type of criteria used by authors in their family therapy Delphi studies.

Most family therapy Delphi studies use purposive sampling, a form of nonprobabilistic sampling, in their research efforts. This form of sampling allows the researcher to choose respondents who best fit study criteria (Bailey, 1982; Goetz & LeCompte, 1984). This type of nonprobabilistic sampling is appropriate for research designs that focus on generalizations to theory, rather than generalizations to populations (Yin, 1989), and fits with qualitative research design (e.g. Goetz & LeCompte, 1984).

Most panelists were identified and recruited through membership lists, personal contacts, and a review of the published works on the researched topic. Some panelists were nominated by other panel-

TABLE 4
Criteria of Experts—Family Therapy Delphi Studies

Criteria of Experts	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*
AAMFT approved supervisors	X					X	X	X		X	X
AFTA members						X	X	X			X
Directors of graduate-level MFT programs	X										
Number of publications on the topic		2	1	1	1				1	1	
Presentations at national conventions		2		1	1				1		
Total years of practice experience		5		5	5				5		
Years of practice with topic									1		
Years teaching and/or supervising topic		5		3	3						
Degree in a mental health profession		X		X	X						
Volunteered to participate	X	X	X	X	X	X	X	X	X	X	X
Knowledgeable about topic*			X								
Editor of family therapy journal									X		

*Others may have considered knowledge important, yet did not separate this aspect as a criterion for experts in the study.

Key for Table:
A = Winkle, Piercy, & Hovestadt (1981)
B = Stone Fish (1985)
C = Watson (1985)
D = Wheeler (1985)
E = Avis (1986)
F = Figley & Nelson (1989)
G = Figley & Nelson (1990)
H = Nelson & Figley (1990)
I = Jenkins (1992)
J = Tester (1992)
K = Nelson, Heilbrun, & Figley (1993)

ists; once identified, they were recruited by the researcher. This type of snowball sampling is deemed helpful in locating and including "nonvisible populations" in a study (Adams & Schvaneveldt, 1991; Taylor & Bogdan, 1984). Although participants volunteered to join the study, this did not ensure full participation throughout the entire project. Follow-up notes and telephone calls were common techniques to encourage completion of the study.

Although each of the reviewed Delphi study researchers whose work is reviewed here attempted to secure a population of knowledgeable panelists for their study, it is unclear how they separated possible panelists who are "more" knowledgeable from those who were "less" knowledgeable. However, Dalkey (1969) suggests that it is not a significant loss to include "less" knowledgeable participants in the study, because they are likely to provide valuable information that may improve the possibility of reaching the desired results.

Panelists' Response Rate

As with most research, securing an appropriate response rate is a constant concern. Authors used follow-up mailings and reminder telephone calls to encourage participation. Figley and Nelson (1989) divided up their questionnaire and mailed only parts of the second questionnaire to participants in order to not overwhelm them with a lengthy questionnaire. Watson (1985) allowed some panelists to participate by telephone, while Tester (1992) permitted members to respond by computer disk in order to facilitate completion of the project. For the studies in which this information is available, percentage rates of the panelists who completed all rounds of the study ranged from 53% to 87%. Table 5 summarizes response rates from panelists employed in each of the Delphi studies.

Item Inclusion Rate

In both qualitative and quantitative investigation, the researcher has the arduous task of managing and making sense of an often immense data set. Due to the overwhelming amount of data that may be gathered in a Delphi study, the ability of panelists to consider every questionnaire item is debatable. Difficulties in reading and considering numerous items in a study affect their eventual inclusion rate in the results of the study. Inclusion rates are defined as the percentage of items that were "acceptable" in the ending profile of the total items solicited in the first questionnaire.

Of the studies that make this information available, inclusion rates of the studies ranged from 17% to 51%. Table 6 illustrates the inclusion rates in the ending profiles for each of the reviewed studies. It is not a question of whether either high or low inclusion rates are more desirable, but whether panelists have the perseverance to carefully consider all items for inclusion in the final profile.

CONCLUSIONS

The Delphi methodology, a combination of both qualitative and quantitative techniques, provides family therapists with a means of refining opinion on a pivotal issue. The areas of design, procedure, and sampling are considered critical to any future research project

TABLE 5
Panelists' Response Rates—Family Therapy Delphi Studies

	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*
Total mailed 1st questnnr	45	46	30	47	39	688	na	na	26	40	na
Compltd 1st questnnr	na	na	30	43	29	429	103	144	21	25	103
Returned 2nd questnnr	na	na	26	38	26	na	na	na	21	21	na
Returned 3rd questnnr	na	32	26	na	na	na	na	na	na	na	na
Number who compltd study	na	32	26	37	26	na	103	144	21	21	103
Initial panelists' response rate	na	70	87	79	67	na	na	na	81	53	na

Note. All numbers, except last row, are totals and not percentages.

Key for Table:
A = Winkle, Piercy, & Hovestadt (1981)
B = Stone Fish (1985)
C = Watson (1985)
D = Wheeler (1985)
E = Avis (1986)
F = Figley & Nelson (1989)
G = Figley & Nelson (1990)
H = Nelson & Figley (1990)
I = Jenkins (1992)
J = Tester (1992)
K = Nelson, Heilbrun, & Figley (1993)
na = not available or not applicable

that incorporates this type of methodology. Careful attention to each of these areas is essential to the effective use of Delphi methodology and to producing solid results.

In Delphi design, choosing the optimal number of rounds to be conducted in the study entails recording the panelists' opinions completely without fatiguing them. Authors continually renegotiate the number of rounds in a study after it has begun, depending on the

TABLE 6
Item Inclusion Rates in the Ending Profile—
Family Therapy Delphi Studies

	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*
# of items in first round	270	213 271	9	301	199	292	na	na	273	328	n a
# of items in endng profle	93	82 97	na	64	62	Top 100	na	na	48	166	n a
% of items in endng profle	34	38 36	na	21	32	na	na	na	17	51	n a

Note. All numbers, besides last row, are totals and not percentages.
*Study included two different surveys to different populations.

Key for Table:
A = Winkle, Piercy, & Hovestadt (1981)
B = Stone Fish (1985)
C = Watson (1985)
D = Wheeler (1985)
E = Avis (1986)
F = Figley & Nelson (1989)
G = Figley & Nelson (1990)
H = Nelson & Figley (1990)
I = Jenkins (1992)
J = Tester (1992)
K = Nelson, Heilbrun, & Figley (1993)

amount of refinement reached early in the study and the fatigue of members.

The design of the questionnaire shapes responses to it. The researcher must carefully consider the previous literature on the topic and decide what questions will best stimulate the thinking of the participants and therefore capture the sought after information. Researchers must consider how their own bias can affect the content of the questionnaire. It is unclear whether the inclusion of preexisting elements from the literature bias, help, or hinder participants in their efforts to refine opinion.

Coding—in which the researcher condenses, summarizes, and

edits responses received from the first round of inquiry into a second questionnaire—is a pivotal step in any Delphi methodology. Although qualitative research suggests that no one understands the data better than the researcher (Goetz & LeCompte, 1984), failure to use independent raters may allow omission and bias in the wording of items in the second questionnaire.

Because there seems to be little agreement on what is the best means to analyze the data received and the literature is not clear what type of analysis best captures the responses of the panelists, researchers should combine both quantitative and qualitative methods. Qualitative measures such as content analysis, voting, and interviews offer information that might otherwise be overlooked in Delphi research.

The sample chosen as "experts" or knowledgeable participants is often the most critical element in any Delphi study. Frequently, researchers attempt to identify all panelists who are acknowledged experts about a given topic. If a sample is needed, researchers need to carefully decide on their sampling technique. Although there are no quantitative guidelines to determine what is an acceptable inclusion rate for a Delphi study, researchers can look to qualitative methods for determination of whether the inquiry is complete.

Standards or guidelines for family therapy Delphi studies have not been carefully articulated. However, qualitative research design, which is often omitted as a guide for Delphi research, offers researchers the needed elements for advancement of this methodology.

Any area of practice that needs additional refinement could benefit from the use of the Delphi method. The greatest benefit from the Delphi method for family therapy at the present appears to be in the emerging areas of practice such as constructivism, language-based (narrative) therapy, and solution-based therapy.

The strength of the Delphi methodologies—the attainment of consensus—also is the greatest weakness. Delphi methodologies encourage the refinement of family therapy practice and theory. In a rapidly growing profession, such work helps establish practice standards, accreditation guidelines, and research priorities. However, consensus can also promote unwanted assertions of practice or theory canon. In the absence of careful study and inquiry, assertions of orthodoxy may stifle creativity and needed advancement in techniques and theory. Thus, results from any one Delphi study should be viewed as a beginning statement and not as a definitive work.

REFERENCES

Adams, G., & Schvaneveldt, J. (1991). *Understanding research methods* (4th ed.). New York: Longmann.

Avis, J. (1986). *Training and supervision in feminist-informed family therapy:A Delphi study*. Unpublished doctoral dissertation, Purdue University, West Lafayette, IN.

Bailey, K. (1982). *Methods of social research* (2nd ed.). New York: The Free Press.

Binning, D., Cochran, S., & & Donatelli, B. (1972). *Delphi panel to explore post-secondary educational needs in the state of New Hampshire*. Manchester, NH: Decision Research, Inc.

Brown, B., Cochran, S., & Dalkey, N. (1969). *The delphi method, II: Structure of experiments*. Santa Monica, CA: Rand Corporation

Cohen, L. & Holliday, M. (1982). *Statistics for social scientists: An introductory text with computer programs in BASIC*. London: Harper & Row.

Dalkey, N. (1969). *The Delphi method: An experimental study of group opinion*. Santa Monica, CA: Rand Corporationn

JDalkeyl'Nn'hqywrin'FIStudies in the quality of life. Lexington, MA: Lexington Books.

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

Dalkey, N., Rourke, D., Lewis, R., & Snyder, D. (1972). *Studies in the quality of life: Delphi and decision making*. Lexington, MA: Lexington Books.

Figley, C., & Nelson, T. (1989). Basic family therapy skills, I: Conceptualization and initial findings. *Journal of Marital and Family Therapy, 15*, 349-365.

Figley, C., & Nelson, T. (1990). Basic family therapy skills, II: Structural family therapy. *Journal of Marital and Family Therapy, 16*, 225-239.

Goetz, J., & LeCompte, M. (1984). *Ethnography and qualitative design in educational research*. Orlando, FL: Academic Press.

Helmer, O. (1967). *The use of the delphi technique in problems of educational innovations*. Santa Monica, CA.: Rand Corporation.

Helmer, O., & Rescher, N. (1959). On the epistemology of the inexact sciences. *Management Science, 6*, 25-52.

Jenkins, D. (1992). *The theory and practice of a reflecting team approach to family therapy: A delphi study*. Unpublished doctoral dissertation. Florida State University, Tallahassee, FL.

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

Martino, J. (1972). *Technological forecasting in decision making*. New York: Elsevier.

Miles, M., & Huberman, A. (1984). *Qualitative data analysis: A sourcebook of new methods*. London: Sage Publications.

Nelson, T., & Figley, C. (1990). Basic family therapy skills, III: Brief and strategic schools of family therapy. *Journal of Family Psychology, 4*, 49-62.

Nelson, T., Heilbrun, G., & Figley, C. (1993). Basic family therapy skills, IV: Transgenerational theories of family therapy. *Journal of Marital and Family Therapy, 19*, 253-266.

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

Redenour, C. (1982). *Future directions in marriage and family therapy in Texas*. Unpublished doctoral dissertation, East Texas State University, Commerce, TX.

Riggs, W. (1983). The delphi technique: An experimental evaluation. *Technological Forecasting and Social Change, 23*, 89-94.

Sackman, H. (1975). *Delphi critique*. Lexington, MA: Lexington Books.

Stone Fish, L. (1985). The theory and practice of structural and strategic family therapies: A delphi study. (Doctoral dissertation, Purdue University). *Dissertation Abstracts International, 46*, 11A.

Stone Fish, L., & Piercy, F. (1987). The theory and practice of structural and strategic family therapies: A delphi study. *Journal of Marital and Family Therapy, 13*, 113-125.

Taylor, S., & Bogdan, R. (1984). *Introduction to qualitative research methods: The search for meanings* (2nd ed.). New York: John Wiley & Sons.

Tester, S. (1992). *The family therapist's use of self: A delphi study.* Unpublished doctoral dissertation. Florida State University, Tallahassee, FL.

Watson, A. (1985). A delphi study of paradox in therapy. In G. Weeks (Ed.), *Promoting change through paradoxical therapy* (pp.). Homewood, IL: Dow Jones-Irwin.

Wheeler, D. (1985). *A feminist model of family therapy: A delphi study.* Unpublished doctoral dissertation, Purdue University, West Lafayette, IN.

Winkle, W., Piercy, F., & Hovestadt, A. (1981). A curriculum for graduate-level marriage and family therapy education. *Journal of Marital and Family Therapy, 7*, 201-210.

Yin, R. (1989). *Case study research: Design and methods* (2nd ed.). Beverly Hills, CA: Sage.