
4 Specific Techniques for Focus Groups, Case Study Research, and Conversational Interviews

More than the others, this chapter (along with Chapters 5 and 6) takes a "how-to" approach in describing techniques for data collection and analysis. This chapter presents specific techniques associated with focus groups, case study research, and conversational interviews (more general issues about the overall design of these methods have been discussed in Chapters 2 and 3). Organizational examples are provided throughout this chapter, and the author recommends that researchers strive to apply multiple techniques within their studies.

Miles and Huberman's influential book *Qualitative Data Analysis* (1994) may be the best single compendium available of specific techniques for qualitative data collection, analysis, and display. Researchers should consult that volume early on in the design or start-up phase of virtually any qualitative study. Whereas a book-length work allows for coverage of a substantial number of techniques, a chapter on techniques requires decisions about what to include and what to omit. Thus, I am omitting here (and in Chapter 5) the techniques associated with grounded theory research, for two reasons. First, organizational

TABLE 4.1 Issues Involved in Focus Group Techniques

Planning focus groups
1. Ethical issues
2. Budget constraints
3. Selection of participants
4. Group structure
5. Group size and discussion quality
6. Number of focus groups
7. Criteria for evaluation
Conducting focus groups
1. Introducing the focus group
2. Beginning discussion
3. Substantive discussion
Data coding and analysis
1. Frequency counts and categorical data analysis
2. Feedback

researchers have treated grounded theory more as a general approach to qualitative research (Larsson & Lowendahl, 1996) than as a particular set of techniques, which was initially the case (Glaser, 1992; Glaser & Strauss, 1967; Locke, 1996; Strauss, 1987). Thus, its application usually involves a subset of the techniques described in Chapters 4 and 5. Second, grounded theory has received substantial coverage in Chapters 2 and 3, so additional coverage here would be redundant. In Chapters 2 and 3, I have discussed the larger issues of research design for focus groups (Morgan, 1997), case study research (Yin, 1994), and conversational interviews (Kvale, 1996); below, I present more technical details related to these topics. More so than other chapters in this volume, this chapter (along with Chapters 5 and 6) takes a "how-to" orientation.

Focus Group Techniques

Planning Focus Groups

Table 4.1 provides an overview of the issues involved in the use of focus group techniques. As is true of many research techniques, the appearance of focus group research can be misleading. To many people, focus groups appear to be simply small collections of individuals who

sit and chat casually about assigned topics. In reality, focus groups are far more complicated. Researchers need to make several important decisions before they can conduct their first focus groups. The issues and decisions that they must address are discussed below. (Much of what follows is derived from the work of Morgan, 1997.)

Ethical issues. Almost always, focus groups are video- or audiotaped. When taping is not feasible (e.g., the host organization prohibits it), the researcher must take extensive field notes on the group discussion. Typically, participants and host organizations are informed that any video or audio recordings made will be used for research purposes only. Interestingly, such vague statements usually satisfy most participants. As a result, perhaps, many academic researchers overlook or do not seriously consider who might eventually see (or hear) these tapes. More often than not, these tapes eventually come to sit on book shelves or in file cabinets. Nonetheless, they often contain material that is potentially embarrassing to participants or the host organizations. Thus, it is strongly recommended that the researcher decide during the planning stage (a) who is likely to see the tapes (or field notes), (b) how long these tapes (or notes) will be kept before they are destroyed, and (c) whether these decisions will be disclosed to potential participants and host organizations prior to an actual focus group.

Budgetary constraints. Of course, all research operates under some budgetary constraints. These tend to be more restrictive for basic academic research, for instance, than for applied market research. It can be critical to a focus group's success for the researcher to decide (or be informed) early on during the planning process how much money is available. For example, should the participants be paid a "market rate" for their time, or should each receive a token sum simply to convey appreciation? It is not uncommon for consulting market researchers to pay each participant in a focus group \$50 to \$100. In basic academic research, however, it is more common to use unpaid volunteers, which is what Maurel, Howe, and Lee (1992) did, or to pay each participant a nominal amount (e.g., \$5 to \$20). In addition, it is often necessary to conduct focus groups at different geographic locations. Traveling to these locations can involve airfare, lodging, car rentals, and meals. Across several locations, expenses can quickly mount.

Selection of participants. As Morgan (1995) notes, recruitment of participants can be quite difficult, if not the most common problem in focus group research. Thus, availability often drives who participates in focus groups. Three recommendations can be offered on the selection of participants. First, to the extent possible, the selection of participants should be based on theoretical sampling (discussed in Chapter 3) rather than random sampling. In other words, researchers should strive to minimize identifiable forms of bias rather than to maximize statistical generalizability. It may be an understatement to say that random selection of participants would be difficult to achieve. Second, and again to the extent possible, participants should be relatively *homogeneous* along a theoretically meaningful dimension (e.g., age, sex, socioeconomic status) but *heterogeneous* across focus groups (e.g., older versus younger, male versus female, higher versus lower status). The minimization of such variability within focus groups often encourages discussion by increasing the likelihood of similarity in perceptions and experiences among participants. The maximization of such variability across focus groups often eases subsequent comparisons and data interpretation. Third, participants who are strangers to one another should be favored over acquaintances. Because acquaintances are more likely to share tacit, taken-for-granted assumptions, discussion among focus group members with stronger acquaintanceship ties can be relatively difficult to interpret, understand, and evaluate.

Group structure. As in the use of the conversational interview, the researcher must make decisions about the focus group's structure. A structured focus group has a preplanned set of agenda topics to be covered, a set of questions to be asked, and a particular desired level of moderator involvement. Typically, a structured format lends itself to discussion of issues that are relatively well understood or for which substantial theory already exists. In contrast, an unstructured focus group has a rather undefined and free-flowing nature. Such a group has only a general topic to guide discussion, questions are not commonly planned in advance, and the moderator strives to manage the group members' interactions as the immediate and changing conditions demand. Within a single focus group session, for example, the moderator may be very directive at times and relatively inactive at others. The unstructured format lends itself to discussion of issues that are poorly

understood or for which minimal theory is available. Between the two extremes of structured and unstructured focus groups (and akin to the semistructured interview), a funnel format might be applied. In the beginning, the focus group might be relatively unstructured, but over time the moderator imposes increasing structure as he or she moves the group along a desired theme.

Group size and discussion quality. Decisions about a focus group's size will directly affect discussion quality. First, the quality of discussion depends heavily upon the participants' involvement with the topic: The greater their involvement, the more likely it is that an active and meaningful discussion will occur. In general, it is easier for members of smaller groups to have higher levels of involvement than it is for members of larger groups. In addition, the moderator should enter with a strong sense of the participants' likely involvement with the topic when the group is small. Second, discussion quality also depends upon the moderator's skill at managing individuals' participation. In a small focus group, for example, an individual's level of discussion can be inhibited by one or two other people. A talented moderator would quickly observe these inhibiting effects and intervene to stop them. In a large focus group, discussion quality can be lost if subgroups begin separate conversations. Again, a skilled moderator would quickly observe the signs that subgroups are forming and intervene to stop them. Thus, a moderator should enter with a realistic sense about his or her facilitation skills.

Number of groups. Although the number of focus groups that can be held is usually affected by cost considerations, ideally the researcher should allow theoretical saturation to determine the number of focus groups used. As noted in Chapter 3, theoretical saturation has been reached when holding one more focus group is not expected to lead to any new learning. More pragmatically, Morgan (1997) recommends that the researcher should conduct at least three focus groups on a given topic. With only a single focus group, the researcher cannot confidently discount the alternative explanation that results are due to the unique combination of participants. Even with the corroboration of a second focus group, confidence in the obtained results still remains somewhat weak. In Morgan's judgment, corroboration from a third group's results indicates that theoretical saturation may be near.

Criteria for evaluation. The criteria to be used to judge a focus group's success should be specified during the planning phase. Four "maximizing" standards for evaluating focus groups have been proposed (Merton, Fiske, & Kendall, 1990; Morgan, 1997). First, the *range* of topics covered during the focus group should be maximized. Operationally, for example, the number of topics, which were either preplanned or emergent, might be counted. Second, the *specificity* of these topics should be maximized. Operationally, for instance, participants' contributions to the discussion could be judged for their concrete and detailed nature. Third, the *depth* of these topics should be maximized. That is, the number of generalities (as opposed to concrete and detailed statements) made by focus group members could be counted as well. Fourth, *differences* in the personal contexts that individuals apply in the interpretation of these topics should be maximized. The number and nature of the personal contexts applied during the members' discussion might also be judged.

Miscellaneous issues. Two additional points merit brief comment. First, focus group members should not feel rushed. They should feel that enough time was scheduled for them to discuss the topic at hand fully. Depending upon the topic, available funding, and the researcher's prior experience, of course, the participants' expectations about the focus group's length might be set for 2 hours. The researcher should be aware that whereas he or she can always bring the session to an end before the allotted time is up if the discussion lags, it will probably not be possible to extend a session past the original expected length, because some (if not most) participants will likely have schedule conflicts. Second, topics should be organized similarly across focus groups. Such similarity can ease the management of group interactions and subsequent data analysis.

Conducting Focus Groups

Perhaps the most important determinant of a focus group's success is the moderator's level of skill in managing group interactions. This skill level is frequently related to the amount of prior experience a moderator has had. In particular, a good deal of prior experience often leads to a moderator's developing an individualized set of tacit knowledge and

skills necessary for conducting successful focus groups. Newcomers to focus group research can best acquire needed moderator skills by observing and participating in focus groups. For experienced moderators, doing what has worked in the past may be the best and most specific advice possible.

Introductions. In a focus group, the members' sense of the discussion's flow, direction, and purpose is critical for success. It is imperative that, at the outset, the moderator set a tone conducive to subsequent discussion. The moderator might begin with candid statements about the overall topic for discussion and then relate his or her goals and expectations for the session. For example, the moderator might state that her intent is to learn what the members think about the topic at hand. By learning from several focus groups, she will not only better appreciate and understand the topic, but might become better positioned to inform others about the topic (e.g., educate managers, create theory). In addition, the moderator should offer general rules for the session; for example, only one person speaks at a time, everyone participates, and no one dominates discussion.

Beginning the discussion. A common tactic used by moderators to open focus group discussion is the icebreaker. For instance, the moderator might ask each participant in turn to tell the group his or her name and give some general biographical background (e.g., hometown, college degrees earned, marital status, names and ages of children) and specific information on his or her personal connection to the focus group's overall topic (e.g., given the topic, how and why are you participating). Sometimes, it may be appropriate to introduce humor; for instance, the moderator might ask each participant to relate the funniest experience he or she ever had at work. (It should be noted that many people are simply not humorous.) Through such preliminary interaction, participants usually become comfortable talking with one another. Ideally, following these opening exchanges, movement through a set of subsequent and preplanned questions occurs as a seamless transition.

Substantive discussion. As may be evident, there are no hard-and-fast rules for initiating, sustaining, and directing conversation. Issues of

timing, sequencing, and probing follow from the unfolding circumstances (as perceived by the moderator). Nonetheless, moderators often use the following three tactics to facilitate conversation. First, a moderator may reiterate participants' points or comments in order to help emphasize key issues that arise during discussion. This can effectively direct conversation toward a particular point or an entirely new topic. Second, a moderator may ask a participant for clarification of particular comments. This helps to freeze conversation on key points, immediately focusing attention and allowing for deeper discussion of those points. Third, a moderator may summarize participants' statements to help move the conversation along. This can serve to maintain (or enhance) the group members' sense of flow and direction.

Miscellaneous issues. Two additional issues should be noted. First, a focus group's physical site must balance the comfort of the participants and the demands on the researcher. For example, a convenient location for participants—such as at a nearby church, university, or community center—may be too expensive for the researcher to rent. Second, as noted above, focus group sessions are usually either audiotaped or videotaped. Often, audio recording is the method of choice because it can typically be accomplished unobtrusively. Microphones and other audio-recording devices easily blend into the background. In contrast, videotaping tends to be very obtrusive. More often than not, people talk and behave differently when they know they are being videotaped; it is difficult to be natural and spontaneous when a camera operator is moving around the room and sometimes pointing the camera directly at you.

Data Coding

There are three common techniques used for coding focus group data, and all three involve counting the number of times a key topic (or issue) is mentioned. The number of times a key topic is mentioned can be recorded according to (a) absolute frequency (the total number of times it is mentioned by any person), (b) frequency by individual persons (the average number of times the topic is mentioned by particular persons), and (c) frequency by individual focus group (the average number of times the topic is mentioned by a group). Moreover, the recorded frequencies might be compared with, for example, a prespecified concep-

tual structure (e.g., a theoretically derived taxonomy). Alternatively, these data may themselves be used to suggest a taxonomic structure.

Data Analysis

There are two tactics for data analysis. First, coding schemes are typically applied, and this often results in frequency counts. Methods of categorical data analysis might be applied (these are described in Chapter 6). Second, a portion of the focus group session might be devoted to the moderator's soliciting the participants' perceptions of what they have learned, or what others might learn, from the discussion. For example, indicators of a topic's importance can be (a) the number of groups that mentioned the topic, (b) the number of people within a group who mentioned the topic, and (c) verbal descriptions about how much energy and enthusiasm the topic generated. These data might be provided (i.e., fed back) to the participants during this designated period, and their interpretations elicited. Often, such interpretations can augment, confirm, or falsify the researcher's current understanding.

In addition, participants might be provided with some quantitative information, such as a rank ordering or rating of the importance of the various topics discussed. Although logistically complicated, a system could be set up whereby the various topics mentioned during the focus group's discussion could be counted as the discussion goes on, then immediately analyzed with categorical data methods (e.g., by a research team member using a laptop personal computer). These statistics could then be fed back to the participants, and their reactions recorded.

Case Study Research Techniques

Although case study methods may be used primarily to describe organizational phenomena, researchers often design their case research along theoretical propositions. Depending upon the specificity of these propositions, they can use their research to test or generate management theories. Thus, the techniques described below can be adapted to either

purpose. In particular, pattern matching and time series methods are discussed. (Much of what follows is based on the work of Yin, 1994.)

Pattern Matching

Although I describe several variations below, all pattern matching involves two simple steps. First, an expected pattern among variables, events, acts, or some other phenomena of interest must be specified before case data are collected. Presumably, the expected pattern derives from a formal theory or set of (less formal) conceptual propositions. Second, the expected pattern is compared with subsequently collected empirical case data for its degree of fit. Note that the logic of pattern matching is no different from that applied to virtually any scientific endeavor. That is, ideas are subjected to empirical test.

Many theories involve multiple outcomes of interest (or dependent variables). Usually, a pattern of increases or decreases in these outcomes is predicted to correspond to changes in a set of theorized antecedent variables. In case study research, the investigator's task is to verify that the changes in antecedent and outcomes variables indeed occurred, but that these changes occurred under conditions set by the theory. Below, a running example is provided, with the job characteristics model as the tested theory. Variations of pattern matching are illustrated using this running example.

The job characteristics model. In their model, Hackman and Oldham (1980) theorize at least five sets of causal linkages. First, increases in a job's collective skill variety, task identity, and task significance enhance the meaningfulness of work. Second, increases in a job's autonomy enhance the experienced responsibility for the work's outcomes. Third, increases in a job's feedback enhance knowledge of the work's results. It should be noted that skill variety, task identity, task significance, autonomy, and feedback constitute the model's independent variables. Fourth, aggregate increases in the (a) meaningfulness of work, (b) responsibility for outcomes, and (c) knowledge of results *increase* an employee's job performance, job satisfaction, and organizational commitment but *decrease* the likelihood of the employee's turnover and absenteeism. Whereas meaningfulness, responsibility, and knowledge of results are the model's intervening variables, performance, satisfaction, commitment, turnover,

and absenteeism constitute the model's dependent variables. Fifth, these linkages are stronger for employees with higher growth need strength and adequate levels of requisite job knowledge and skills. Thus, growth need strength and job knowledge and skills constitute the model's contextual (or moderator) variables.

Measurements. In case study research, each of the model's variables, linkages, and conditions might be measured as follows. First, analysis of engineering and purchase documents might verify whether *objective* changes to the company's computer system actually increased (or decreased) a job's levels of skill variety, task identity, task significance, autonomy, and knowledge of results. Second, employee interviews or participant observation could corroborate whether the theorized increases (or decreases) in employees' experienced meaningfulness, responsibility, and knowledge of results occurred. Third, analysis of individual personnel records or employee interviews could indicate whether job performance, job satisfaction, and organizational commitment increased and whether turnover and absenteeism decreased. Fourth, inspection of preemployment testing records might reveal individuals' levels of growth need strength and requisite knowledge, skills, and abilities.

Matching dependent variables. Presuming variables were validly measured and specified intervening and contextual conditions occur as theorized, pattern matching of the dependent variables involves judging whether these outcomes changed as theorized. For the job characteristics model, pattern matching of the dependent variable involves judging whether job performance, job satisfaction, and organizational commitment actually went up and whether turnover and absenteeism actually went down. If the predicted pattern is observed, corroboration of the model should be inferred. If the predicted pattern is not observed, falsification of the model should be inferred. If some portion of the expected pattern is observed, some corroboration or falsification can be inferred, depending upon one's preference of terms.

Matching independent variables. Pattern matching of the independent variables involves specification of one or more (if possible) mutually exclusive combinations of the independent variables and the subsequent

observation of which combination of variables actually occurred. Under ideal circumstances, observation of one particular combination of dependent variables would preclude the others. For instance, the job characteristics model specifies that a *multiplicative combination* of skill variety, task identity, and task significance affects experienced responsibility. Subsequent observation that a linear, geometric, or otherwise noncompensatory combination of these components increases (or decreases) experienced meaningfulness constitutes falsification of this portion of the model. Similarly, subsequent observation that the specified multiplicative combination affected experienced responsibility or knowledge of results also constitutes falsification for this portion of the job characteristics model. (Autonomy is theorized to increase felt responsibility, and feedback is theorized to increase knowledge of results.)

Replication of matching dependent or independent variables. Although informative, single case studies must be viewed with caution. As in any study, replication is recommended. If a second organization holds an identical or a very similar set of a theory's variables and linkages as that found in the original case, a *literal* replication can be conducted. If, on the other hand, the second organization holds a different set of a theory's variables and linkages than that found in the first case, a more informative *theoretical* replication may be possible. Suppose, for example, the variables and linkages for the job characteristics model were validity measured in a second organization. Suppose further that the second organization's level of employee need growth strength was dramatically low, to the point of being virtual nonexistent. A theoretical prediction would be *no* (or minimal) change in the dependent variables. If a substantial change is observed in the dependent variables, theory falsification should be inferred, and replication did not occur. If no (or minimal) change is judged to have occurred, theory corroboration should be inferred, and theoretical replication may have occurred. If a minimal change is judged, some corroboration or falsification can be inferred, and a partial replication might be inferred as well.

Time Series

Much like time series in quasi-experimental design (Cook & Campbell, 1979), time series can be conducted in case study research. A strength

of case study research is that organizational phenomena and the local contexts in which they occur can be followed (prospectively) over time or reconstructed (retrospectively). One or more dynamic patterns for organizational events, acts, circumstances, or variables can be predicted; data can be collected (prospectively or retrospectively); and a comparison can be made between the predicted and empirical patterns. Alternatively, an expected pattern need not be predicted in advance. Instead, an empirical strategy could be followed, whereby the phenomena of interest are simply monitored or reconstructed.

Simple designs. With simple time series designs, the pattern of a minimal number of variables (e.g., one) and the contexts in which they occur are monitored over time. In the simplest design, a variable (e.g., employee's job performance) might be monitored over time in order to establish a base rate. On the one hand, observation of a slightly positive trend may be sufficient to inhibit managerial interventions, because such actions can be quite expensive or disruptive. On the other hand, a slightly negative trend may be sufficient to encourage managerial actions, because of concern for long-term consequences.

Slightly more complicated designs. In a slightly more complicated design, an *interrupted time series* might be applied. Sometimes, an intervention can be anticipated, such as the introduction of a new technology, compensation package, or organizational structure. If a base rate has been established during the preintervention period, changes in a monitored variable (e.g., job performance over time) before and after the intervention could be compared (judgmentally or statistically). Changes from the base rate after the intervention might be reasonably attributed to the intervention itself.

Still more complicated designs. In a still more complicated design, company or government records might allow for a historical reconstruction of a phenomenon's base rate across multiple operating units (e.g., output at a consulting firm's 10 branches located in different cities). Prospectively or retrospectively, the introduction of an intervention might be documented, and the levels of the phenomenon can be compared between pre- and postintervention periods across the multiple case replications. Changes from the base rate after the intervention might be

reasonably attributed to the intervention itself. From these three simple variations, the strength of a causal inference is enhanced.

Complex designs. With more complex designs, changes in trends can be monitored or reconstructed over time. For example, the phenomenon of interest might be changes in output levels across three manufacturing plants after the introduction of team-based management. It might be expected, for example, that, in comparison to a base rate, output falls somewhat after the immediate implementation of teams because of ambiguities in the new system. After a year, it might be expected that output will increase substantially. With historical records of plant output, supplemented with interviews of people who worked in these three plants during the period under study, a predicted decrease followed by an increase in output could be compared with what actually happened. In addition to trend changes on a single variable, multiple variables could be monitored for single or multiple trends. For example, employee morale, turnover, and acts of sabotage may be of concern among those who survive three rounds of layoffs at 10 company plants. Because layoffs can often be anticipated well in advance, a prospective case design could be adopted. In particular, the three outcomes of interest could be monitored and base rates established. Based on various survivor theories (e.g., Noer, 1993), it might be predicted that survivors' morale and turnover would decrease and that acts of sabotage would increase. Immediately after each round of layoffs, the predicted pattern of the outcomes might be compared with their (a) base rate levels and (b) levels immediately after the prior layoff experience. Again, changes from the base rate after the intervention might be reasonably attributed to the intervention itself.

Chronologies. Like other forms of "more complex designs," chronological case studies also follow multiple variables and changes in trends over time. In addition, one or more of the following conditions are imposed onto data collection, analysis, interpretation, and display:

1. "Some events must always occur before other events, with the reverse sequence being impossible."
2. "Some events must always be followed by other events, on a contingency basis."

3. "Some events can only follow other events after a prespecified *passage of time*."
4. "Certain *time periods* in a case study may be marked by classes of events that differ substantially from those of other time periods." (Yin, 1994, p. 117)

To the extent that one or more of these conditions can be met, it is often asserted that causal inferences are substantially strengthened.

Comment. Yin (1994) also delineates "lesser modes of analysis," for which he recommends survey data and other cross-case comparisons. In essence, Yin advocates "counting the countable" (Cassell & Symon, 1994). He cautions, however that such quantitative analyses may be difficult because of the inherently small sample sizes in case study research. Nonetheless, methods of categorical data analyses can and should be employed when feasible (e.g., Lee, Mitchell, Wise, & Fireman, 1996).

Conversational Interview Techniques

Planning the Conversation

A major characteristic of the conversational interview is its openness. To those who are unfamiliar with this methodology, a conversational interview may appear to be simply two individuals casually chatting about whatever comes to mind. In practice, the conversational interview is really much more than that. Although there are no hard-and-fast rules for the conduct of such interviews, researchers need to address several issues before they undertake to use this methodology. Below, I discuss the issues of thematizing, design, situation, criteria for evaluation (i.e., validity, which is also discussed in Chapter 6), and transcription as they relate to conversational interviews (see Table 4.2 for an overview of these issues). (Much of what follows is based on the work of Kvale, 1996.)

Thematizing. Because conversational interviews can be quite unstructured (appearing aimless to some observers), the researcher should enter with a high degree of conceptual clarity as to the interrelationships

TABLE 4.2 Issues Involved in Conversational Interviews

Planning the conversational interview
1. Thematizing the interview
2. Designing the interview
3. Focusing the situation
4. The interview questions
5. Judging the interview's quality
From speech to text
1. Audiotaping
2. Videotaping
3. Field notes
4. Working from memory
Transcriptions
1. Reliability
2. Validity
Major modes of data analysis
1. Meaning condensation
2. Meaning categorization
3. Narrative structuring
4. Hermeneutic meaning interpretation
5. Ad hoc methods
Implications
1. Triangulate with multiple modes of data analysis
2. Create hybrid modes of data analysis

among the study's purpose, the research question, and the analytic method. In other words, the researcher should be able to articulate the conversational interview's themes as he or she begins the planning process. This thematizing, as it is sometimes called by qualitative researchers, might include reference to existing theory, literature, and the researcher's insight. Common themes include, for example, whether the conversational interview should be (a) exploratory or confirmatory, (b) relatively unstructured versus relatively structured, (c) intended to generate or test theory, and (d) the study's primary or secondary data collection method. To the extent possible, the researcher should do this thematizing in writing, and then share it with peers for informal review.

Comment. It might be noted that thematizing is sometimes too quickly overlooked by those qualitative researchers who hold a strong antipositivist orientation (e.g., some postmodernists). These researchers often

want to begin directly with the design phase (discussed below). In particular, they frequently reject the value of prior and conceptual planning because it can be incompatible with a strong emphasis on reflexive research designs (i.e., sufficient flexibility to respond to local conditions). In other words, they prefer themes to emerge from the data and designs to be maximally contextualized (or highly sensitive to local conditions). Although certainly not wrong, these views may overly restrict the conversational interview technique to exploratory and theory-generating purposes. In my judgment, such limitations to the conversational interview minimize its practical value. Thus, I strongly encourage researchers to do as much thematizing (i.e., prior planning) as possible.

Design. Whereas *thematizing* refers to the articulation of what a study should do and why it should do so, *design* refers to the overall plan of how to conduct the study. It includes such technical issues as (a) the number of interviews to be conducted, (b) whether the interviews should include more factual versus more interpretative questions, and (c) whether the study's various parts lend themselves to (if not facilitate) the final outcome (e.g., a 40-page double-spaced journal submission). The acts of thematizing and design should clarify for the researcher in a compelling fashion when the conversational interview is well- versus ill-suited to the study in question. For example, Kvale (1996) suggests that conversational interviews should be avoided if thematizing or design reveals a researcher's interest to be in (a) estimating a population's parameter value, (b) the direct study of behavior, or (c) deeper understanding of organizational phenomena that operate at a tacit level.

Focusing the interview situation. The conversational interview involves two (or more) persons chatting on themes of mutual interest. However, the persons involved have an unequal power relationship. Almost always, the interviewer is the driving force. Thus, the interviewer's framing of the interview (how he or she defines the situation) for the interviewee becomes critical to success. Framing involves (a) explaining to the interviewee the interview's purpose (e.g., research study); (b) providing an overview of the interview agenda (i.e., topics and questions), the expected nature of responses, the extent of informality, and the probing nature of some follow-up questions; (c) stating that a debriefing

will be provided at the interview's end; and (d) warning that field notes, tape recordings, or both will be made during the interview.

The interview questions. Kvale (1996, pp. 133-135) suggests that there are nine interrelated types of interview questions (listed below). Although these are presented in an apparent sequence, the informal and relatively unstructured nature of the conversational interview requires that the interviewer move back and forth among the various types of questions throughout the interview session, as the situation's dynamics suggest. Again, substantial interviewer skill, sensitivity, and insight are demanded.

1. *Introductory questions* elicit personal experiences about the phenomena understudy and frame subsequent questions. Typical introductory questions include the following: "Can you tell me about . . . ?"; "What happened when . . . ?"; and "Have you ever experienced . . . ?"
2. *Follow-up questions* serve, ideally, as seamless transitions to new issues. These questions should be asked in a sincere, curious, persistent, and critical manner, which can be enhanced by such interviewer behaviors as head nods, pauses, and prodding utterances (e.g., "yes"). In addition, the interviewer might notice and ask about the interviewee's unusual word choices, body language, and changes in intonation.
3. *Probing questions* are used to seek new information without necessarily stating the theme the interviewer wishes to pursue. Common probing questions include the following: "Could you say more about . . . ?"; "Can you give another example to help clarify . . . ?"; and "What did you mean by . . . ?"
4. *Specifying questions* move the discussion toward greater specificity. These questions commonly follow more general questions or interviewee responses. For example, the interviewer might ask: "What did you actually do at that moment?"; "Did you have a physical, as well as emotional, reaction?"; and "Has this, or anything like this, ever happened before?"
5. *Direct questions* introduce new topics and themes. Because of their directness, they can be disruptive or jarring to the ongoing conversation. Interviewees can feel threatened and become inhibited during subsequent conversation. Thus, the interviewer must show considerable care and sensitivity in asking direct questions such as, "Have you ever stolen anything worth more than \$10 dollars from work?" or "When dealing with your former boss, did you ever behave toward her in a mean-spirited fashion?" Given the potential sensitivity of direct questions, interviewees must feel that at least a modest degree of openness and trust has been established.

6. *Indirect questions* concern specific topics but allow for either specific or general responses. Interviewees may respond personally or in reference to a general group. For instance, "What are the typical reasons people quit their jobs here?" can be answered from the viewpoint of a general work group, a person who recently left, or the respondent's own likely behavior. Similarly, "What do you believe is the general morale around here?" can be interpreted and answered from multiple perspectives.
7. *Structured questions* are typically prewritten items that are used to shift conversation to a new topic when a current topic is unlikely to generate any more discussion. In approaching such a question, the interviewer might say, "I'd like to move to a new topic."
8. *Interpretative questions* involve restating, rephrasing, summarizing, or paraphrasing the interviewee's remarks. These questions often serve to facilitate conversation, redirect discussion, and avoid misunderstanding. The interviewer may begin, for example: "Let me see if I understand you—did you mean to say . . . ?"; "If I heard you correctly, you're saying . . . ?"; or "Let me summarize what you've said. . . ."
9. *Silence* is not technically a type of question, of course, but well-placed pauses (or "screaming gaps of silence"), augmented by the interviewer's expectant facial expressions, can often be quite effective at eliciting comments.

Interview quality. A high-quality conversational interview is one that earns affirmative answers to the following questions:

1. Was the conversation spontaneous, rich, and specific?
2. Were answers relevant to the questions asked?
3. Were the interviewer's questions short and the subject's responses long?
4. Did the interviewer follow up or clarify the meanings of the interviewee's answers throughout the interview?
5. Did the interviewer interpret interviewee responses throughout the interview?
6. Did the interviewer corroborate interpretations during the course of the interview?
7. Did the interview itself appear self-communicating?

High-quality work on the part of the interviewer is reflected by affirmative answers to the following questions:

1. Was the interviewer knowledgeable about the topic of the interview? That is, did he or she know what themes or aspects to pursue, and did he or

- she determine that the interviewee had substantial background information about the organization and work context?
2. Did the interviewer impose sufficient structure on the interview? That is, did he or she make the experience appear neither too routinized nor too free-flowing?
 3. Did the interviewer ask clear, understandable, and (usually) short questions? For instance, did the interviewee seldom need to ask for clarifications or restatements?
 4. Was the interviewer sufficiently gentle or sensitive to the issue at hand? In other words, did he or she show empathy or tolerance for more ambiguous, less well articulated responses?
 5. Was the interviewer a good listener? In particular, did he or she wait for the interviewee to complete responses, show interest in the responses, and respond quickly to new issues?
 6. Was the interviewer critical? That is, because responses should not necessarily be taken at face value, did he or she conduct interpretation of the responses on an ongoing basis?
 7. Was the interviewer able to remember and refer back to earlier discussion? That is, was he or she able to connect points and issues raised throughout the interview?

From Speech to Text

Conversational interviews are recorded most often through some combination of audio- or videotape, field notes, and simple memory; a combination of methods is useful because each of these recording methods has its strengths and weaknesses.

Audiotaping, videotaping, and field notes. Audiotape captures an exact record of the spoken conversation, but it does not record visual aspects of the physical context, facial expressions, or body language. Typically, however, the interviewer takes notice of these aspects of the interview and records such characteristics and their potential effects in extensive field notes. In contrast, videotape yields an exact record of the spoken conversation as well as much of the physical context and some of the nonverbal communication. However, the process of videotaping is quite obtrusive and can itself produce unintended effects. In comparison with "natural conversations," people often behave and verbalize differently while they are being videotaped. In addition, videotaping can produce a huge amount of data, attempted analysis of which may well over-

whelm the researcher's personal and financial resources. Thus, the decision to videotape should be made with care. Whether the researcher decides to employ audio- or videotape, it is highly recommended that he or she also take extensive field notes to both supplement and enhance the interpretation of the mechanical tapes.

Comment. As we approach the new millennium, we often take product reliability for granted; however, mechanical failures still sometimes occur. Because conversational interviews are time-consuming for both parties and potentially difficult to reschedule, the interviewer must take steps to avoid mechanical failure. First, it is a good idea for the interviewer to bring along extra batteries, an extra tape recorder, or both. Second, the interviewer should not assume that he or she is familiar with the particular recording device to be used (e.g., how it might connect to a telephone). Further, the interviewer should be aware that such easily overlooked elements as poor placement of a microphone, a particular room's acoustics, and an interviewee's low, monotone voice can result in an inaudible recording.

Another comment. Although it is not advisable for an interviewer to try to rely on memory alone, on occasion this may become necessary. Once, my colleagues and I began a study with the strong general expectation that everyone we interviewed would be comfortable being audiotaped. We routinely asked for permission, but we never expected anyone to refuse it. We were quite startled and unprepared when one interviewee, who had volunteered to participate, requested no audiotaping or manual note taking. Although we wrote copious field notes immediately after the interview, the quality of those notes were suspect because of our lack of preparation. Eventually, we dropped the data from this interview from the study because of concerns over the data's reliability. Thus, as I have noted, although sole reliance on memory is ill-advised, my own experience suggests that a researcher should be prepared for the (fortunately few) situations when this may become necessary.

Transcripts

The recordings of conversational interviews are rarely analyzed directly. In addition to augmenting their field notes, researchers almost

always have interviews transcribed. The reliability and validity of these transcriptions must be considered. Simply put, unreliable or invalid transcription renders a conversational interview useless.

Reliability. Ordinarily, a variation of interrater agreement is applied to estimate transcription reliability. For example, two persons type up separately an entire interview (or selected portions of an interview) using a computerized qualitative research software package. The program then identifies, locates in the file itself, and counts word differences between the two typed versions. (Many common word processing packages have this capability as well.) A percentage of correct overlap could be calculated as an estimate of reliability. An adequate, less dependable, but cheaper approach would be for one person to transcribe the text and another to compare the printed text with what he or she hears on the tape.

Comment. The estimation of reliability must be done. It is conceptually straightforward, mechanically simple, and compatible with the norms and values of the larger discipline of organizational science. More pragmatically, journal reviewers might (and should) require some meaningful estimate of reliability. A word of caution is in order, however. Because estimating transcript reliability can be so straightforward and mechanically simple, it is easy to underestimate the skills required of the typists. In a resource-rich environment, the use of professional transcribers can resolve many concerns. In a resource-poor environment, it is often convenient and cost-efficient to employ an undergraduate, master's, or doctoral student to do this work. I have used both professional and student transcribers in my own work, and I have found that whereas quality has been consistently high among the professionals, it has been extraordinarily variable across student transcribers. With respect to transcription, then, an old adage seems to apply: You get what you pay for.

Validity. In comparison to reliability, transcription validity is a far more difficult issue. Potentially, a transcription can fundamentally alter the original discourse's meaning. In no small way, the change from oral to written language often entails application of different rules. From the viewpoint of a strong adherent of postmodernism, for example, transcription directly decontextualizes and detemporalizes the text's mean-

ing. To such a postmodernist, the text's meaning is thus so fundamentally altered that it is unlikely to be reflective of the actual interview. In other words, although reliable, the transcript is an invalid representation of the interview's content. To a more mainstream quantitative researcher, who may hold a more positivist view, a transcription does not necessarily alter a text's meaning in such a dramatic fashion. This latter researcher would likely perceive much more room for judging how much alteration has occurred. That is, there is room for an assessment of whether too much, insufficient, minimal, or no change in meaning has been made.

Comment. Because I discuss methods of estimating and judging validity in Chapter 7, further detail will be deferred. At this point, however, some researchers, particularly those who are very strong adherents of postmodernism, social construction, and emergent meanings, might pose the question of validity as, What is the correct transcription? Under those conditions and to those persons, the question may be impossible to answer. To the vast majority of "mainstream" qualitatively and quantitatively oriented management researchers, however, the question of validity might be posed differently. These persons might choose to ask, What is a useful transcription for research? In my judgment, the second question is more valuable than the first.

Major Modes of Data Analysis

I discuss below five modes of analyzing the interview data: meaning condensation, meaning categorization, narrative structuring, hermeneutic meaning interpretation, and ad hoc methods. It is important to note that each mode is typically individualized by different researchers. That is, each researcher tends to impose his or her own preferences and style in operationalizing the specific steps involved in each of these modes of data analysis.

Meaning Condensation

During meaning condensation, the researcher strives to extract, abridge, or abstract the most important themes from the data (i.e., the interview's text). As the name suggests, meaning condensation involves data reduction, while simultaneously articulating the data's most impor-

tant themes. Meaning condensation can be seen as analogous (albeit somewhat loosely interpreted) to a statistical factor analysis, where the "latent roots" or underlying causal constructs from a correlation (or covariance) matrix are identified and extracted.

Meaning condensation typically involves five basic steps. First, one or more members of the research team read the entire transcript of an interview. (Some researchers prefer to complete the entire five-step analysis one interview at a time; others prefer to read across an entire set of interviews before beginning the second step. My personal preference is to do both—that is, I like to read across all interviews in order to develop a sense of the entire data set, and then analyze each conversational interview, one at a time.)

Second, the researcher identifies "natural meaning units," or portions of the text that are judged to relate to an identifiable theme (or issue). These natural meaning units can consist of sentence fragments, complete sentences, portions of paragraphs, or longer passages. The researcher pulls out these portions of text and reassembles them physically into a continuous flow of text. Before word processing packages made manipulation of portions of text easy, researchers would make photocopies of the text, cut out potentially connected portions, and array them on a tabletop or paste them onto other sheets of paper. The researchers would then make decisions as to these units' naturalness, holism, and relevance to an identifiable topic (or theme). Now, virtually any word processing package or software written for qualitative data analysis makes the gathering of potential natural units quite simple. Making decisions about their naturalness, holism, and relevance remains, however, a creative and challenging act of induction.

In the creation of natural meaning units, conceptual completeness, thoroughness of content, and substantive depth should be driving forces. Nonetheless, it is also advisable for the researcher to have some sense of how large a quantity of data is manageable. On the one hand, if too little text is extracted, the natural meaning unit may contain an insufficient amount of information and be uninterpretable. On the other hand, if too much text is extracted, the natural meaning unit may contain an overwhelming amount of information and also be uninterpretable. Returning to the earlier analogy to factor analysis, natural meaning units can be seen as akin to principle factors (or extracted latent traits).

Third, the natural meaning unit must be clearly defined. Sometimes this step is called thematizing because the unit's underlying theme (or latent trait) is to be conveyed. Often, a label and an accompanying short paragraph can effectively capture and convey the natural unit's substantive content. Like its creation, writing the natural meaning unit's definition is also a creative and challenging act. Unlike its creation, defining the unit is a deductive exercise.

Yin (1994) describes the fourth step as "interrogating the meaning units in terms of the specific purpose of the study" (p. 194). In other words, the researcher asks how each natural meaning unit fits with or informs the study's research questions. These meaning units are judged for how well they concisely capture and convey a major idea from the data.

Fifth, the concisely worded natural meaning units must be integrated into a coherent and nonredundant structure. That is, the larger data set must be reduced into a more meaningful, coherent, yet manageable set of underlying themes. Returning to the analogy to factor analysis, the outcome of meaning condensation is akin to the achievement of a factorial "simple structure."

Meaning Categorization

Whereas meaning condensation reduces the data to their most important themes, meaning categorization distributes statements from the interview data into quantifiable categories. These categories can be binary, such as "It occurs"/"It does not occur," or they can be ordinal, such as a rating of from 1 to 5 indicating an increasing magnitude for "it." The object of these categories (or the "it") can be a job attitude, work behavior, organizational characteristic, or other managerial phenomenon. The outcome of these categorizations can be analyzed qualitatively or quantitatively.

Meaning categorization usually involves four steps. First, the researcher must select an organizing conceptual structure that will serve as the basis for subsequent categorizing judgments. That is, the structure defines the phenomena and variables of interest. Moreover, it can be derived from an existing theory, model, or set of propositions from the literature, or it can be derived from the interview data themselves (in an iterative fashion with subsequent steps).

Second, the researcher must define as clearly as possible the focal attitudes, behaviors, organizational characteristics, and other phenomena of interest identified by the conceptual organizing structure. That is, the researcher must fully specify the characteristics or dimensions that will be evaluated subsequently. Ideally, these characteristics or dimensions are independent of one another. Less ideally, though certainly feasible, these characteristics (or dimensions) and their subdimensions might be modestly correlated (or dependent). The stronger this correlation, however, the more difficult will be the subsequent categorizations, analyses, and interpretations.

Third, one or (preferably) more evaluators must rate each interview transcription on each defined dimension. If, for example, there are 13 dimensions, each interviewee will have 13 scores. Presuming multiple evaluators, each dimension's binary or ordinal scores could be compared for interrater reliability.

Fourth, the researcher may analyze these scores qualitatively through the use of various graphic depictions (e.g., tables, charts, or cross-tabulations; see Miles & Huberman, 1994, for a compendium of display options). Alternatively, he or she may analyze the scores quantitatively using categorical data analysis (described in Chapter 6) or more traditional methods (e.g., factor analysis, discriminant analysis, or regression).

Narrative Structuring

Whereas meaning condensation and meaning categorization seek to simplify data in a relatively straightforward manner, narrative structuring seeks to identify and reconstruct the interview text into longer stories. From responses to open-ended questions (e.g., Can you tell me, please, about the situation and circumstances surrounding the last time you quit your job?) or across several sequential questions (e.g., [a] Was there a particular event that first caused you to think about leaving? [b] Was that event expected or unexpected, positive or negative, or work or nonwork related? [c] Did that event lead directly to quitting or did that event lead to a careful comparison of the pluses and minuses of quitting?), a continuing narrative can sometimes be detected. Through the extraction and rearrangement of relevant text, a more continuous, coherent, integrative, and engaging single story can be recovered. Such

a recovered narrative can often stimulate insight into the explanations for an organizational phenomenon of interest.

The work of narrative restructuring involves three steps. First, the researcher reads the transcribed interview as a whole. Second, he or she undertakes subsequent readings to identify temporal sequences, recurring social dimensions, and any overall plot in the text. Third, the researcher goes through an interactive (i.e., trial-and-error) process of physically arranging and rearranging the text itself, using a conventional word processing package or qualitative data analysis software.

Ultimately, the ideal outcome would be one or more compelling narrations. Each narration should be complete, with a beginning, a story or plot, and an ending, and each narration would be presented as a "stand-alone" story.

Hermeneutic Meaning Interpretation

In meaning condensation, meaning categorization, and narrative structuring, the researcher's interpretation is drawn almost exclusively from the interview text itself. In contrast, hermeneutic meaning interpretation requires that the researcher *impose* meaning based on the perspectives from a preexisting paradigm. For example, interview texts might be interpreted from such well-known perspectives as psychoanalysis, literary criticism, feminism, or the philosophy of Karl Marx or Ayn Rand. Because the mechanics of hermeneutic meaning interpretation vary according to the requirements of the adopted paradigm, specific steps and outcomes cannot be readily prescribed in advance. However, it is quite common for such interpretations to vary greatly in length; some may be longer than the original interview transcriptions themselves (e.g., a literary critique of a poem).

Ad Hoc Methods

In the four modes described above, the logic and mechanical steps involved in data analysis can be more or less anticipated. As the name suggests, ad hoc methods involve the application of any technique, based on either personal preference or trial and error. Miles and Huberman (1994, pp. 245-246), for example, describe 13 tactics that can be

useful in the ad hoc interpretation of texts. Going from primarily descriptive to explanatory (Kvale, 1996), these tactics are as follows:

1. Noting patterns and themes
2. Seeing plausibility
3. Clustering
4. Applying metaphors
5. Counting
6. Comparing and contrasting
7. Partitioning variables
8. Subsuming particulars under general categories
9. Factoring
10. Noting relations between variables
11. Finding intervening variables
12. Building a logical chain of evidence
13. Creating theoretical coherence (i.e., theory building)

Conclusion

Although the techniques described in this chapter are associated with focus groups, case study research, and conversational interviews, it should not be overlooked that application of *multiple* techniques is highly advised. In other words, no method is perfect, and it is only prudent to triangulate with multiple methods. In addition to the techniques described here, the methods described in Chapter 5 (and the quantitative methods described in Chapter 6 as well) can often be quite appropriate for use in organizational research. Thus, researchers may want to design some of the techniques described in Chapters 5 and 6 into their studies (e.g., in two-phase or mixed-methodology designs).