Examples of Mixed Model Designs

As we discussed in Chapter 3, mixed model Types I through IV are different from Patton's (1990) prototype designs in that instead of differentiating designs as experimental versus naturalistic, we differentiate them according to their degree of being confirmatory (including qualitative case studies, experimental designs, and nonexperimental studies) versus exploratory (including naturalistic inquiry as well as quantitative exploratory studies such as surveys). Some examples of these designs are presented below and also in the next chapter.

Please note that the Patton examples presented in Box 7.1 are specific cases of our mixed model design Types I, II, III, and IV, as follows:

Type I: Confirmatory investigation, qualitative data, statistical analysis (form 2 in Box 7.1)
Type II: Confirmatory investigation, qualitative data, qualitative analysis (form 1 in Box 7.1)
Type III: Exploratory investigation, quantitative data, statistical analysis (form 4 in Box 7.1)
Type IV: Exploratory investigation, qualitative data, statistical analysis (form 3 in Box 7.1)

Patton did not include examples of Types V and VI designs because they are rare, involving qualitizing techniques (see Chapter 6).

CONFIRMATORY INVESTIGATIONS

The mixed model confirmatory designs were noted in Table 3.1 as mixed model designs Types I, II, and V. historically, in confirmatory studies, the data primarily have been quantitative, the conceptual framework has been deductive, and the data analysis has been statistical. However, in the mixed confirmatory designs discussed here, the data can be qualitative or quantitative and can be analyzed in either form as well. At least tentative
BOX 7.1
Examples of "Mixed Forms" of Evaluation Research From Patton

These are possible projects to evaluate a program serving high-risk students who were likely to be involved with the criminal justice system.

(1) Mixed form: Experimental design, qualitative data, and content analysis. As in the pure experimental form, potential participants are randomly assigned to treatment and control groups. In-depth interviews are conducted with all youth, both those in the treatment group and those in the control group, before the program begins. The focus of those interviews is similar to that in the pure qualitative approach. Interviews are conducted again at the end of the program. Content analysis is performed separately on the data from the control group and the experimental group. The patterns found in the control group and the experimental group are then compared and contrasted.

(2) Mixed form: Experimental design, qualitative data, and statistical analysis. Participants are randomly assigned to treatment and control groups, and in-depth interviews are conducted both before the program and at the end of the program. These interview data, in raw form, are then given to a panel of judges who rate each interview along several outcome dimensions operationalized as a ten-point scale.

For both the "pre" interview and the "post" interview, the judges assign ratings on such dimensions as likelihood of success in school (low = 1, high = 10), likelihood of committing criminal offenses (low = 1, high = 10), commitment to education, commitment to engaging in productive work, self-esteem, and manifestation of desired nutritional and health habits. Inferential statistics are then used to compare these two groups. Judges make the ratings without knowledge of which participants were in which group. Outcomes on the rated scales are also statistically related to background characteristics of participants.

(3) Mixed form: Naturalistic inquiry, qualitative data, statistical analysis. As in the pure qualitative form, students are selected for the program on the basis of whatever criteria staff choose to apply. In-depth interviews are conducted with all students before the program and at the end of the program. These data are then submitted to a panel of judges, who rate them on a series of dimensions similar to those listed in the previous example. Change scores are computed for each individual, and changes are statistically related to background characteristics of the students to determine in a regression format which characteristics of students are likely to predict success in the program. In addition, observations of program activities are rated on a set of scales developed to quantify the climate attributes of activities: for example, the extent to which the activity involved active or passive participation, the extent to which student-teacher interaction was high or low, the extent to which interactions were formal or informal, and the extent to which participants had input into program activities. Ratings of activities based on qualitative descriptions are then aggregated to provide an overview of the treatment environment of the program.

(4) Mixed form: Naturalistic inquiry, quantitative data, statistical analysis. Students are selected for the program according to staff criteria. The evaluator enters the program setting without any predetermined categories of analysis or presuppositions about important variables or variable relationships. The evaluator observes important activities and events in the program, looking for the types of behaviors and interactions that will emerge. For each new type of behavior or interaction, the evaluator creates a category and then uses a time and space sampling design to count the frequency with which those categories of behavior and interaction are exhibited. The frequency of the manifestation of observed behaviors and interactions is then statistically related to such characteristics as group size, duration of the activity, staff-student ratios, and social/physical density.

NOTE: These are possible projects to evaluate a program serving high-risk students who were likely to be involved with the criminal justice system.

predictions are made and tested in investigations using these designs. The following sections address the three types of designs in this group.

Type I: Confirmatory investigation, qualitative data and operations, statistical analysis and inference

In this type of study, the collected data are qualitative; they are quantitized and then subjected to statistical analysis. This is one of the most frequently used mixed designs in the literature. The Pemberton, Insko, and Schopler (1996) study of group versus individual competitiveness is an example of this design. An initial experiment (see Pemberton et al., 1996, p. 954, for details) had demonstrated that college students recalled more competitive intergroup interactions than interindividual interactions. Following that experiment, the authors asked questions regarding the reasons that intergroup interaction was perceived as more competitive. Tentative answers were suggested on the basis of theory and previous research:

(a) awareness of experimental expectations (discussed in Chapters 4 and 5 (continued)
under participant reactivity or participant roles), (b) existence of an outgroup schema, and (c) memory effects (previous group experiences have been in sports, where competition is the norm).

In the Type I study that followed, Pemberton and his colleagues used the Rochester Interaction Record (RIR; Reis & Wheeler, 1991) data collection procedures to record interactions in the natural environment. A number of hypotheses/predictions were formulated and tested; for example, it was predicted that participants would recall more competitive group events. Also, it was predicted that observation of actual events recorded through the RIR would show greater competitiveness in intergroup interactions than in intragroup ones.

The participants were 28 male and 27 female introductory psychology students. As a part of the RIR procedures, each participant was given a booklet containing 40 sheets, one sheet for recording each interaction. Each participant was asked to record the dates and circle the type of interaction that occurred (e.g., one-on-one, one-on-group, group-on-one, and group-on-group).

An interesting characteristic of this experiment was that the participants were used as raters for quantizing the qualitative data (interactions). Each recorded event was rated on a 7-point scale at the time of recording (as soon as the participants recorded the event). The ratings pertained to the perceived degree of competitiveness or cooperativeness of the interaction. These quantized data were then statistically analyzed and reported.

The Smith, Sells, and Clevenger (1994) study is another example of this type of design. During an investigation on reflecting team practice in marriage counseling, they collected qualitative data, quantitized parts of them, and analyzed them quantitatively using chi-square and the phi correlation coefficient. Box 7.2 presents a summary of that study.

Type II: Confirmatory investigation, qualitative data and operations, qualitative analysis and inference

This mixed design is basically a confirmatory qualitative study. In contrast to the traditional naturalistic design, these investigations use tentative predictions in at least some aspects of the research. The most common studies in this group are the ones that use qualitative data collection and analysis to confirm or disconfirm (triangulate) the findings from a previous study.

Sinclair's (1994) study is an example of an investigation in which this type of design was the dominant mode. Sinclair used qualitative data and procedures to test four hypotheses regarding the effects of prediction making as an instructional method in science classrooms. Participating students were divided into two groups: an experimental group that used prediction-making activities as part of a genetics unit and a control group that was taught a genetics unit in a traditional manner. Qualitative data collection involved student interviews, teacher interviews, teacher logs, and classroom observations. Qualitative analysis consisted of the use of Bogdan and Biklen's (1982) constant comparative method in which responses were put into categories.

Qualitative analysis showed that prediction activities led to greater student involvement and interest in class. Students in the experimental

**BOX 7.2**

An Example of Type I Mixed Model Study

The Smith et al. (1994) study was a confirmatory investigation involving unstructured interviews with couples who participated in marital therapy sessions. In a previous study, they had collected qualitative data consisting of interviews of couples in therapy and their therapists. Qualitative analysis of the data led to two themes. First, the therapists and the couples' perceptions differed with regard to reflecting team practice. Second, these differences were associated "with the spatial distance or boundary" between the couples and the "client-index therapist system" (p. 269). The purpose of the current study was to confirm hypotheses based on these two assertions.

Qualitative data were collected through ethnographic interviews of couples and therapists during a four-month period. The categories of the previous study were used in the initial qualitative analysis of the data. Content analysis led to the refinement of seven categories. Frequency of occurrence for each of these categories of themes/ perceptions was recorded and quantitatively analyzed.

For the first hypothesis, chi-square was used to compare the frequency of each category between the couples' and the therapists' perceptions. Results showed that both groups discussed gender with the same frequency. Although the therapists were most concerned with the use of reflecting team practice in the sessions, the couples were most concerned with the benefits of the practice. Phi correlation between the co-occurrence of "Spatial Separateness" and "Process of Hearing" was calculated to test the second hypothesis. A strong correlation was found between these two. "These co-occurrences of spacial separateness and process of hearing were supported by many examples throughout the text" (p. 280).
group participated more in class, asked more “thoughtful” questions, and seemed to enjoy the classes more than those in the control group. The observers reported greater “give-and-take” dialogues between the students and the teachers in these classes than in the control group. Qualitative analysis also revealed that differences in the teaching styles of the teachers affected the way the classes in both groups were conducted, and probably affected the results of the study.

Another example of this type of design is a study by Mann (1994) in which qualitative data were collected regarding two groups of tutors. One group consisted of three tutors who seemed to have benefitted the most from a training program (they had the highest scores on outcome measures). The control group consisted of those tutors who had not benefitted from the training program (had low scores on the outcome measures). The general hypothesis was that the two groups of tutors would be different in their conceptualization of their role as well as their effects on the tutee.

Qualitative data collection involved weekly journal entries during a 10-week period and a two-page essay written by each tutor. Qualitative analysis of the journal entries showed that the tutors with high scores were more aware of the interpersonal aspects of tutoring, identified conflicts in clearer forms, and resolved them more efficiently than the low-scoring group. The low-score tutors were more preoccupied with evaluating the sessions and with overgeneralizing the problems of the tutee. They experienced more anxiety and less success than the high-score tutors. (For detailed verbal examples of these journal entries, see Mann, 1994.)

Type V: Confirmatory investigation, quantitative data and operations, qualitative analysis and inference

Although this type of design is rare, examples are sprinkled throughout the literature in several fields of study. The quantitative data are qualititized and presented/analyzed qualitatively.

Typical Type V studies identify and present verbal “profiles” for groups and/or individuals on the basis of quantitative data. An example of this kind of study might be an investigation designed to identify the attributes of different groups or “types” of female contraceptive users in an urban center in Zaire on the basis of interview (survey) data. The typology is developed on the basis of previous research (for example, consistent user, sporadic user) but a holistic profile (see Chapter 6) is developed for each type or group on the basis of survey data. See examples of this type of design in Chapter 8.

EXPLORATORY STUDIES

The mixed model exploratory designs were noted in Table 3.1 as mixed model designs Types III, IV, and VI. These mixed model exploratory studies might use qualitative or quantitative data and analyze them using either of the two approaches. The following sections present three different types of designs in this group.

Type III: Exploratory investigation, quantitative data and operations, statistical analysis and inference

This type of design is actually the same as traditional quantitative exploratory investigations (usually known as descriptive studies) in which no predictions were made beforehand. The reason for inclusion of these studies as mixed model studies is that they do not fit in the traditional hypothetico-deductive type of investigation using the logical schemata that we presented in Table 3.1. This parallel between traditional descriptive and mixed Type III designs is a good example of a phenomenon that will be expanded upon in Chapter 9: Researchers working in traditional areas (e.g., psychology, anthropology, sociology) may be using mixed model designs but may be so ingrained in the QUAN or the QUAL tradition that they believe they are doing “pure” QUAN or QUAL research.

A recent exploratory study by Aghajianian and Moghad (in press) is an example of a Type III design. They investigated recent divorce trends in Iran, including potential determinants and consequences of these patterns. As is common in exploratory studies, they had no a priori hypotheses. The data were collected through a survey study in a large urban center. A multilevel sample of married and divorced women was selected on the basis of the 1986 census. In the first stage, 43 census districts were randomly selected from 169. All households that had at least one divorced woman were identified in the census data. The result was a sample of 254 women who had been married once and divorced once. A comparison (2% random) sample was also selected consisting of women who were currently married and never divorced. Female interviewers visited each household and interviewed the identified woman. Detailed data were collected regarding socioeconomic characteristics, attitudes, health, and problems.

The results suggested that the divorce rate has evolved in response to social and legal changes and several years of war with Iraq. The main differentiators between the two groups of women (divorced, not divorced)
were found to be urbanicity, education, work status, and religiosity. A greater proportion of divorced women were from urban centers, had either low or high education (as compared with a middle level), and were less religious. Regarding the consequences of divorce, divorced women were found to suffer economically and experience more psychological problems than nondivorced women. Also, the children of divorced women experienced higher levels of emotional problems and delinquency than those of nondivorced women. It was concluded than the consequences of divorce in Iran are much more severe than those found in the United States. Also, the inverse relationship between the age of first marriage and probability of divorce that is reported in the United States was not found in Iran.

Type IV: Exploratory investigation, qualitative data and operations, statistical analysis and inference

A large number of studies have used projective techniques (sentence completion, stories, and so on) for collecting data in psychological and sociological research. After collection, the data are quantitized (see Chapter 6) through different procedures and analyzed through (mostly nonparametric) techniques such as log-linear modeling and logistic regression.

Type VI: Exploratory investigation, quantitative data and operations, qualitative analysis and inference

This is another relatively rare design in which quantitative data are collected and subjected to qualitative analysis and presentation. The discussion of quantitzing the quantitative data presented in Chapter 6 included examples of this type of design. In such studies, groups of participants are formed (e.g., extreme-case groups, types of participants) on the basis of the quantitative information. These groups are then described in qualitative (narrative) form (e.g., what we called “profile formation” in Chapter 6).

The Rusbult et al. (1993) study (discussed in detail in Chapter 8) is an example of Type VI mixed model studies. In this exploratory study, participants were asked to rate a set of descriptions regarding romantic relationships. These descriptions were rated with regard to their degree of similarity to a set of “target” descriptions. Based on multidimensional scaling analysis (a multivariate statistical analysis), four types of relationships were identified in these descriptions. These four prototypes were then described in detail on the basis of both qualitative and quantitative information. Box 7.3 presents two examples of these prototypes (profiles).

**BOX 7.3**

Examples of the Different Prototypes (Profiles) in the Rusbult et al. Study

*Picturebook fantasy (female)*: “Quadrant 1 mental models received low ratings for Intimacy and high ratings for Romance-Traditionalism and were highly idealized, developed quickly and spontaneously, included explicit requirements for involvement, were not egalitarian (e.g., traditional sex roles were common), de-emphasized friendship, were not based on respect/admiration, and tended to be male initiated. Rather than seeking trust and intimacy, these subjects yearned for the classic romantic fantasy—glamorous activities, an exciting life, and passion. Love at first sight was a common theme. These ideals are termed Picturebook Fantasy and are illustrated by the following excerpts: It would be love at first sight. My partner would have an endless number of good qualities. He would never pressure me sexually—he would be very mature about it, unlike most guys. . . . I would be in love almost to the point of obsession, and I would never before have felt so happy.

*Companionship (male)*: “Quadrant 4 mental models were the mirror image of those in Quadrant 1 and were characterized by slow development, egalitarianism, friendship orientation, a practical orientation to living with the partner, high respect/admiration, the absence of male initiation, and few explicit requirements for the relationship. This is a very comfortable, low-key ideal—partners are buddies and allies, sharing career interests and political beliefs and life concerns. The following excerpts illustrate Quadrant 4 Companionship: It would revolve around trust and mutuality. We’d be “up front” with each other—I don’t enjoy playing games. Honesty would be very important. Also, I like a girl who is aggressive, I don’t like to do all the work and take all the risks. . . . She’d be able to talk to other men and I wouldn’t feel jealous or “inadequate.” She’d love me for what I am, not for what I have or who I associate with or what fraternity I belong to.


Another example of this type of design is a study by Taylor and Tashakkori (1997) discussed in Chapter 8. The authors collected survey data from a sample of teachers in a restructured school district. The researchers had no a priori hypotheses. Two specific types of questions were included in these surveys: regarding teachers’ desire/motivation for participation in decision making and their reports of actual involvement in decisions regarding instructional and school issues. Four extreme groups
of teachers were formed on the basis of their scores on these two dimensions (desire for participation versus actual participation). The result was the identification of four "types" of teachers in this two-dimensional model. These four types were labeled "empowered" (those who wanted to participate and did), "disenfranchised" (those who wanted to participate but did not), "involved" (those who did not want to participate but were involved in decision making), and "disengaged" (those who did not want to participate and did not). Verbal (qualitative) modal profiles were constructed on the basis of other information that was available about each group of teachers.

8

Extended Examples of Mixed Model Designs

As we discussed before, completely mixed investigations simultaneously use both types of data collection (qualitative and quantitative) and both types of data analysis (statistical and qualitative analysis). Some may also combine the exploratory approach with a confirmatory research study. Box 8.1 presents a graphic illustration of mixed model strategies that were used in a large-scale international study of women's reproductive health (Ulin et al., 1996). As the illustration shows, the degree and type of mixing may vary across subprojects of this large-scale study.

We present three extended examples of completely mixed investigations in this chapter. The ultimate goal of the chapter is to provide graduate students and interested researchers with more concrete examples of completely mixed studies that can be used as models.

In Table 1.1, we refered to these types of studies as mixed model studies with multiple applications within each stage of research (type of inquiry, data collection/operations, or analysis/inference). A note to the table further explained that mixing must occur such that both approaches appear in at least one stage of the study. The examples in this chapter have mixing at both the data collection/operation and the analysis/inference stages.

The first extended example (Taylor & Tashakkori, 1997; Taylor, Tashakkori, & Hardwick, 1996) is a parallel mixed model study: It combined QUAL and QUAN data collection, data analysis, and an inference process in a parallel form (multiple simultaneous approach). The inference process drifts simultaneously between inductive and deductive, and the data collection and analysis are simultaneously qualitative and quantitative.

The second example (Rusbult et al., 1993) is different in that each phase of the investigation was dependent on the finding, conceptual development, or material that was generated in a previous phase. As such, that study can be more readily considered to be an example of sequential mixed model investigations. In these studies, multiple approaches to data collection,
analysis, and inference are employed in a sequence of phases. Each phase, by itself, may use a mixed approach and provide conceptual and/or methodological grounds for the next one in the chain.
Table 8.2
Five Phases of the Louisiana School Effectiveness Study

<table>
<thead>
<tr>
<th>Phase</th>
<th>Brief Description</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Conducted pilot study in 1981-1982, including field test of instruments</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Completed LSES-I report</td>
<td></td>
</tr>
<tr>
<td>LSES-II Macro-level</td>
<td>• Selected sample of 76 schools</td>
<td>1982-1984</td>
</tr>
<tr>
<td>(process-product study)</td>
<td>• Collected school climate questionnaires and other instruments from 74 principals,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>250 teachers, and more than 5,400 students in 1982-1983</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Analyzed data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• LSES-II report completed in 1984</td>
<td></td>
</tr>
<tr>
<td>LSES-III Micro-level</td>
<td>• Selected sample of eight matched pairs of schools</td>
<td>1984-1988</td>
</tr>
<tr>
<td>longitudinal study</td>
<td>• Conducted more than 700 hours of classroom observations and more than 1,000 hours</td>
<td></td>
</tr>
<tr>
<td>(case studies, first-round</td>
<td>of on-site data collection in 1984-1985</td>
<td></td>
</tr>
<tr>
<td>visits)</td>
<td>• Analyzed data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Prepared research reports and articles using LSES-III quantitative and qualitative data</td>
<td></td>
</tr>
<tr>
<td>longitudinal study</td>
<td>• Conducted more than 650 hours of classroom observations and more than 950 total</td>
<td></td>
</tr>
<tr>
<td>(case studies, second-round</td>
<td>hours of on-site data collection in 1989-1990</td>
<td></td>
</tr>
<tr>
<td>visits)</td>
<td>• Analyzed data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Prepared research reports and articles using both LSES-III and IV data sources</td>
<td></td>
</tr>
<tr>
<td>LSES-V Micro-level</td>
<td>• Published Schools Make a Difference: Lessons Learned from a 10-Year Study of School</td>
<td>1992-1996</td>
</tr>
<tr>
<td>longitudinal study</td>
<td>Effects, summarizing LSES-I through LSES-IV in 1993</td>
<td></td>
</tr>
<tr>
<td>(case studies, third-round</td>
<td>• Designed LSES-V during 1994-1995</td>
<td></td>
</tr>
<tr>
<td>visits)</td>
<td>• Conducted more than 500 hours of classroom observations and more than 750 total</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Began data analysis</td>
<td></td>
</tr>
</tbody>
</table>

An investigation of school restructuring effects by Taylor et al. (1996; Taylor & Tashakkori, 1997) combined qualitative and quantitative data collection, analysis, and inference processes in the form of a completely mixed study. In that study, all teachers in a school district were asked to respond to a survey instrument with closed-ended questions. The instrument combined selected subscales from the National Educational Longitudinal Study of 1988 with a subscale from an instrument developed by Bacharach, Bauer, and Shedel (1986). Items asked for teachers' perceptions regarding school climate, student discipline, teachers' sense of efficacy, and teachers' participation in decision making. Response choices were arrayed in a four-part Likert format ranging from strongly disagree to strongly agree. An open-ended section was added to collect qualitative responses by asking teachers to share their feelings and opinions about their school and related issues.

In addition to collecting survey data, classroom observations in each of the seven schools were made by trained observers. Observations occurred in at least eight classrooms per school on two different occasions. Extensive field notes recorded not only data from these classroom observations but also details about other aspects of each school. A third component of the research design involved focus groups that were conducted separately with elementary and middle/high school teachers, elementary and middle/high school students, school administrators, school board members, and parents.

Data analysis consisted of both quantitative and qualitative techniques. Univariate and multivariate analyses were conducted on the survey data.
Content analysis was performed on the open-ended responses, field notes, and transcribed focus group interviews. As part of these analyses, some of the qualitative data were quantified by counting the frequency of occurrence of events. Also, some of the quantitative data were qualitatively by forming groups of teachers on the basis of responses to the survey instruments (see the example in Chapter 7). Brief profiles were constructed for each of these groups. Results were presented both in terms of quantitative indices and in terms of qualitative inferences.

AN EXAMPLE OF A SEQUENTIAL MIXED MODEL STUDY

Although the investigators did not identify their methodology as “mixed,” a study by Rusult et al. (1993) is an interesting example of such a design. It is an interesting example for at least three reasons: (a) the authors efficiently combined experimental procedures with qualitative data collection and inference; (b) they conducted a set of sequential studies, each using the material and/or conceptual framework obtained in the previous phase; and (c) the study is an example of a research project that could have benefited from the use of detailed and systematic qualitative content analysis of the qualitative data, such as discussed in Chapter 6.

The study was based on an initial review of the literature, including the social-psychological theories of liking, love, and romantic relationships. The main research question, following these reviews, was as follows: “Do people hold expectations about ideal romantic involvement that are compatible with these theories?” (p. 498). Not finding an answer to this general question in the social-psychological literature, four specific questions were raised to guide the study: (a) What dimensions distinguish among young adults’ mental models of ideal involvement (i.e., how do mental models of ideal involvement differ)? (b) What are the primary mental models of ideal involvement? (c) Do stated ideals parallel the ways in which young adults think about actual or potential relationships? (d) Are there substantive differences between men’s and women’s mental models of ideal involvement?

The investigation started in a descriptive, qualitative, and largely inductive direction. However, as new studies (phases) were added to the investigation, the researchers collected mostly quantitative data and analyzed them statistically. Three studies were conducted to explore the “young adults’ mental models of ideal romantic involvement” (p. 493). Each of the first two studies was conducted in three phases.

Phase 1 of the project involved qualitative data, partial qualitative analysis, and development of materials for Phase 2 of the study. Participants (undergraduate students) were asked to describe ideal romantic involvement (qualitative data). The obtained descriptions were edited for grammar and clarity, and each was printed on a small index card. The ratings of multiple judges regarding each expression were solicited and summed to obtain an overall index of the degree of clarity.

From the set of descriptions that were judged to be the most clear (overall clarity score above the median), 50 (25 expressed by men and 25 by women) were selected for the next phase (in a second study, all 35 descriptions were retained). Ten of the descriptions were selected (and were called “targets” in the rest of the report). The participants were asked to judge the degree of similarity of all other descriptors to these 10 “targets.”

Phase 2 of the study also involved having qualitative data quantified by the respondents and analyzed statistically. The respondents were asked to rank order the qualitative descriptors obtained in Phase 1 according to their degree of similarity to a “target” scenario that was assigned to each participant. They were specifically asked to pay attention to the ideal of romantic involvement that was expressed in the description. After ranking the descriptions, each participant was also asked to write a brief summary of the criteria he or she used to distinguish among the descriptions (either these qualitative data were not analyzed, or their qualitative analysis is not reported). Difference scores were calculated between the ranks assigned to each possible pair of descriptions. These “dissimilarity” scores were then subjected to multidimensional scaling (ALSCAL; see Takane, Young, & DeLeeuw, 1977).

Based on different statistical considerations, a two-dimensional model (configuration) was identified to distinguish the difference between the descriptions. One of the two dimensions consisted of intimate versus superficial relationship. The other consisted of a romantic/traditional versus practical/nontraditional type of relationship. Based on this (2 x 2) configuration, four categories of ideal relationships were identified: picturebook fantasy, marital bliss, utilitarian involvement, and companionship.

The two dimensions of the “mental model” that represent the participants’ thoughts about relationships, and the four quadrants that are obtained in such a two-dimensional model, could not be labeled on the basis of statistical analysis. Phase 3 of the investigation was an attempt to obtain qualitative interpretations and/or labels for these dimensions and quad-
rants. From Phase 2, 39 potential labels were developed for different configurations of relationships. These labels were obtained from several sources: participant's descriptions in Phase 2, investigators' interpretations of the obtained statistical dimensions, and models and theories of romantic love in the social-psychological literature.

The study (in multiple phases) identified the types of romantic relationships and their possible differences. The next research question asked by Rusbult et al. was whether the identified mental models were the result of daydreams or were indeed the same as the actual mental representations that young adults use in evaluating their romantic relationships. A new study was designed to answer this question. Participants in this study were 25 men and 26 women who self-identified themselves as being involved in a dating relationship for at least a month. They were asked to write essays in response to one of three detailed instructional sets describing their (a) ideal romantic relationship, (b) what they look for in a romantic relationship, and (c) their current romantic relationship.

Judges were trained to rate each description with regard to its correspondence to the category labels obtained in Phase 3 of the previous study. For each category, ratings assigned to descriptions were averaged to obtain a single category score. These category scores were then compared across the three scenarios (instructional sets; see above).

Analysis of variance revealed very few differences between the three settings. Also, category labels were divided into four groups corresponding to the four quadrants obtained from earlier (Phase 3) statistical analyses. Correlational analyses were performed to assess the degree of correspondence between the category scores in the four obtained quadrants. Based on these analyses, the authors concluded that the "mental models" that were obtained in the previous study were consistent with the real (actual) ones obtained in the current study.

A STEP-BY-STEP EXAMPLE OF SIMPLE AND COMPLEX MIXED MODEL DESIGNS IN A SINGLE STUDY

The Louisiana School Effectiveness Study was conceived in 1980, and final data collection occurred in spring 1996. During its 16+-year history, the study has progressed through five phases of data collection, interspersed with periods of analysis and report writing. The five phases are summarized in Table 8.2, and the following section will briefly describe the overall design of the study.

The Design of LSES-I and LSES-II

The first and second phases of the LSES (LSES-I and LSES-II) were conducted primarily as confirmatory studies, while the final three phases (LSES-III through LSES-V) included both exploratory and confirmatory components. The LSES was always driven by basic research questions. Specific and innovative research methodologies evolved for answering these questions during the history of the project. The following section on the overall designs of LSES-I through LSES-V was adapted from Teddle and Stringfield (1993).

The pilot study (LSES-I) was conducted during school year 1981-1982. Two separate activities were accomplished during LSES-I: (a) The school climate questionnaires, modified from those used by Brookover et al. (1979), were field tested in one school district and refined based on the information gained there; and (b) the entire methodology for LSES-II was pilot tested in a number of schools in a second district. In the field test of the modified student school climate questionnaire, the authors first surveyed third grade students in regular classroom settings and then interviewed small groups of students to determine if they understood the questions and to gather information on how the instrument might be improved.

Data for LSES-II were collected during the 1982-1983 school year (Teddle, Falkowski, Stringfield, Desselle, & Garvue, 1984). The sampling frame for LSES-II was designed to make the study sample representative of the statewide population of elementary schools. Altogether, 76 schools with more than 250 third grade teachers and more than 5,400 third grade students were included in the sample. Twelve school districts participated in LSES-II, including urban, suburban, and rural areas.

The major instruments used in LSES-II were a norm-referenced achievement test (NRT) and the modified Brookover et al. (1979) student, teacher, and principal school climate questionnaires. The student questionnaires were mass administered during regularly scheduled classes, and the principals and teachers completed the instruments on their own. Teams of researchers entered each of the 76 schools and gathered the data in one day. Archived data (e.g., student SES data, results from state-administered criterion-referenced tests, or CRTs, and school structural characteristics) were also collected from several secondary sources.
LSES-II was conceptualized as a large-scale process-product study similar to that conducted by Brookover et al. (1979) in Michigan. The emphases in data analysis were twofold: (a) to determine the amount of variance in student achievement that was attributable to student SES and to measures of school climate and (b) to compare schools that varied in terms of effectiveness status and SES characteristics of student body.

The first set of analyses used both multiple regression and HLM (hierarchical linear modeling) techniques. The second set of analyses (multivariate and univariate analyses of variance or MANOVAs/ANOVAs) enabled the researchers to make the following comparisons: (a) differences among effective, typical, and ineffective schools; (b) differences between middle- and low-SES schools; and (c) differences among the six groups of schools generated by the design.

The Design of LSES-III Through LSES-V

These phases were designed to determine the degree to which individual schools retained their effectiveness status over time despite changes in student body and staffing. These phases were designed to provide longitudinal case studies of eight matched pairs of effective and ineffective schools during an 11-year time period. Schools for these three phases were purposefully selected to represent different geographic regions and urbanicity contexts within the state. The sampling procedure was complex, involving nine separate steps (Teddlie & Stringfield, 1993). Data collection occurred in 1984-1985 (LSES III). The same schools were revisited in 1989-1990 and 1995-1996 during LSES-V.

Numerous instruments were used in the study. They included student, teacher, and principal school climate instruments; NRT achievement tests for all third grade students; protocols for interviewing teachers and principals; a 12-page school observation checklist; high-inference classroom observation instruments (measuring quality of teaching); low-inference classroom observation instruments (measuring time-on-task); an instrument designed to assess school-level climate that was completed by the researchers; and sociometric indices to assess relationships among faculty members. These instruments included a mix of QUAL and QUAN measures.

Observers spent several days in each school during these three phases of the study (12 person days per school in LSES-III, 10 person days in LSES-IV, 10 days in LSES-V). A major focus of LSES-III through LSES-V was the gathering of classroom observational data on teachers in matched pairs of effective and ineffective schools. Therefore, more than 700 hours of classroom observations were recorded in LSES-III, some 650 hours were logged in classrooms in LSES-IV, and 500 hours were spent in the classrooms in LSES-V. The results from these classroom observations constituted a large portion of the results from this part of the study.

In the following sections, we present examples from the LSES of all eight of the pure and mixed research designs described in Table 3.1.

Purely Quantitative Study

LSES-II followed in the tradition of quantitatively oriented educational production function studies conducted in school effectiveness research. One of its major purposes was to determine the proportion of variance in student achievement that was attributable either to family background (e.g., SES of students attending the school, other student body demographic characteristics) or to the influence of the school (e.g., school climate). This portion of the LSES was purely quantitative. The study was confirmatory, with hypotheses set up in advance asserting that the school climate variables would predict a significant proportion of the variance in student achievement, thus proving that “schools make a difference.” The data collection and operations were purely quantitative, with the use of paper-and-pencil attitudinal scales, NRT data, and numerical archived data. The analyses and inferences were purely quantitative, using regression or HLM statistical techniques and interpreting the results in light of the study hypotheses and previous research results.

Purely Qualitative Study

The longitudinal case studies (LSES-III through V) were set up with exploratory research questions and confirmatory hypotheses, employed both qualitative and quantitative data and procedures, and used both qualitative data analyses and statistical analyses. Therefore, in general, these case studies are examples of what we call "completely mixed investigations" or "multiple applications within stage of study" mixed models in Table 1.1.

While the case studies were completely mixed investigations, there were independent studies within them that were purely qualitative in nature. The best example of a purely qualitative design within the LSES was the "naturally occurring school improvement" study from LSES-III reported by Stringfield and Teddlie (1990). In this study, the researchers “discovered” a phenomenon that had not been discussed in school effectiveness research: Some schools previously designated and confirmed as ineffective
schools were improving, without any external change agent being involved. The initial data collection and operations for this portion of the LSES were purely qualitative, based on narrative data obtained through field notes, interviews, observations, and audiotapes. The analyses and inferences were purely qualitative, using the inductive reasoning process solely and the constant comparative technique in particular. Themes related to the process of "naturally occurring school improvement" emerged as a result of these qualitative analyses.

Type I Mixed Study: Confirmatory Investigation, Qualitative Data/Operations, Statistical Analysis and Inference

The research reported by Teddlie et al. (1989) is an example from LSES-III of a Type I mixed design in which the authors of a confirmatory study used the quantifying methodology described in Chapter 6. Open-ended results from the Classroom Observation Instrument (COI) were converted to numerical data, which were then analyzed using t-tests and MANOVAS/ANOVAS. The researchers set out to determine if different types of classroom teaching behavior occurred in differentially effective schools. They predicted that more effective schools would be characterized by teachers who provided higher quality instruction than teachers in less effective schools. Teacher quality was measured by the Stallings’s Classroom Snapshot (Stallings, 1980; Stallings & Kaskowitz, 1974), which resulted in numerical indices of time-on-task, and the COI.

The COI was developed to provide high-inference classroom data in the form of field notes regarding 15 general indicators of teacher effectiveness. These indicators were based on characteristics of effective teaching gleaned from reviews by Rosenshine (1983) and Rosenshine and Stevens (1986) and included measures of time-on-task, initial student practice, presentation of new material, and positive reinforcement. Specific cues associated with "initial student practice" included a high frequency of questions, teacher-directed exchange, teacher prompts, opportunity for all students to respond, and success rates of 80% during initial student practice. Observers were directed to complete field notes in each of the 15 categories as behavior associated with that category occurred in the classroom being observed.

During LSES-III, observers coded teaching and student behaviors in 700 separate classes. Altogether, more than 25,000 units of information were recorded (700 classes x 15 observation items on the COI x an average of 2.5 observer comments per item). To reduce the overwhelming quantity of data, they were quantified. COI field notes were analyzed by independent raters. Considering all notes for any given teacher, two raters scored each of the indicators of effective teaching from 1 (strong evidence of effective teaching) to 2 (contradictory evidence) to 3 (effective teaching not evident). The results of the statistical analyses (t-tests and MANOVA/ANOVA) indicated that teachers in more effective schools did indeed demonstrate much higher quality teaching than those in less effective schools.

This portion of the LSES was a Type I mixed model design. The study was confirmatory, with hypotheses set up in advance predicting that more effective schools would be characterized by teachers who provided higher quality instruction than teachers in less effective schools. The data collection and operations were purely qualitative, with observer responses to the COI generating a large amount of narrative data in the form of field notes. The analyses were quantitative, using standard statistical analyses on data that had been converted to numbers using quantifying techniques.

Type II Mixed Design: Confirmatory Investigation, Qualitative Data/Operations, Qualitative Analysis and Inference

Another aspect of the Teddlie et al. (1989) research involved the use of the case study approach to further understand the processes whereby more effective teaching occurs in more effective schools and less effective teaching occurs in less effective schools. A finding of the study was that the teachers in more effective schools displayed more consistent behaviors (i.e., the standard deviation of their behaviors was smaller) than teachers from less effective schools. One purpose for examining the qualitative data was to determine how these differences in the behavior of teachers (i.e., both the mean and the variance differences) occurred.

The LSES-III qualitative data consisted of the field notes taken while observing throughout the school and in the classrooms as well as the interviews conducted with the principal and the teachers. The researchers selected one matched pair of schools in which the differences in quality of teaching between the two schools was quite large. The researchers then examined the field notes from these two schools closely to try to determine the differential ongoing processes. Their qualitative descriptions were included in the results section of Teddlie et al. (1989) and focused on the principals at the differentially effective schools: One was a proactive "guardian" of academic time at her school, while the other allowed academic time to "slip through the fingers of the faculty."

This portion of the LSES was a confirmatory study, with hypotheses set up in advance predicting that more effective schools would be charac-
terized by teachers who provided higher quality instruction than teachers in less effective schools. The data collection and operations were purely qualitative, consisting of the field notes taken while observing throughout the school and in the classrooms as well as the interviews conducted with the principal and the teachers. The analyses and inferences were qualitative, using inductively oriented methods in which the field notes were examined for emerging themes that would describe the processes underlying the statistical differences that had previously been uncovered.

Type III Mixed Design: Exploratory Investigation, Quantitative Data/Operations, Statistical Analysis and Inference

The Stringfield and Teddlie (1990) study, described above as using a purely qualitative design, also employed a Type III mixed design at the insistence of the editors of School Effectiveness and School Improvement, who required that the researchers look at quantitative data and analyses before they would publish the purely qualitative results. These editors wanted to be sure that the investigators were observing “real” school improvement, and they insisted that the researchers provide “hard” data to that effect (i.e., standardized test data indicating that the schools were scoring better than they had done in the past).

Thus, ex post facto, the researchers examined CRT and NRT data for the schools, using information from the 1982-1984, 1984-1985, and 1989-1990 school years. The results from these quantitative data were congruent with the previously gathered qualitative data, confirming that some type of improvement was occurring in the four schools under study. This portion of the LSES was exploratory, given that no hypotheses regarding “naturally improving” school improvement had been set up in advance. The data collection and operations were quantitative, involving researcher-administered NRT and state-mandated CRT data. The analyses were statistical in nature, involving comparisons of data generated from the four improving schools with their previous test results and with results from their matched pairs.

Type IV Mixed Design: Exploratory Investigation, Qualitative Data/Operations, Statistical Analysis and Inference

Crone and Teddlie (1995) conducted a partial replication of previous research (e.g., Teddlie et al., 1989; Virgilio, Teddlie, & Oescher, 1991) regarding teacher behavior in differentially effective schools. In the interview component of this study, researchers queried teachers and principals regarding reasons for the different patterns of results found in teacher behavior in effective and ineffective schools: higher rates of effective teaching behavior and less variance in that behavior in more effective schools, and lower rates of effective teaching behavior and more variance in that behavior in less effective schools.

The researchers used an “interview guide approach” (Patton, 1990) in which topics and issues were specified in outline form in advance, which allowed the interviewers to determine the exact sequence and wording of the questions while the interview was ongoing. The topics were rather broad, concerning the selection and socialization of teachers, the cohesiveness of the faculties, the principals’ implementation of change, and so on.

The researchers specified no hypotheses in advance for the interview component of the study, which was viewed as exploratory, designed to determine if teachers and principals in differentially effective schools viewed their experiences (e.g., selection and socialization) differently. The data gathered were narrative field notes taken in response to the open-ended questions from the interview guide.

As the researchers began to analyze the data using the constant comparative technique described by Lincoln and Guba (1985), a number of themes began to emerge. These themes appeared to be occurring with more or less frequency depending upon whether the interviewers were working in effective or ineffective schools. The researchers decided to count the different themes that emerged and then subject that numerical data to chi-square analyses.

This analysis technique is another example of the quantizing methodology described in Chapter 6. Analysis of the interviews showed that there was more cohesiveness among the teachers in effective schools. . . . This included 13 teachers in effective schools and 14 teachers in ineffective schools. . . . All teachers were asked what kind of input they had regarding school goals. Twelve of the teachers in effective schools responded that teachers worked together. Only two teachers in ineffective schools answered that teachers worked together on goals (Chi-square = 7.17, p < .01). (Crone & Teddlie, 1995, p. 5)

This component of the Crone and Teddlie (1995) study was a Type IV mixed model design. The study was exploratory, with no hypotheses set up in advance; the data collection and operations were qualitative, using interviews that resulted in narrative field notes; and the analyses and inferences were quantitative, employing field notes to analyze frequency count data that had been quantitized from the narrative data.
Type V Mixed Design: Confirmatory Investigation, Quantitative Data/Operations, Qualitative Analysis and Inference

Types V and VI mixed model designs are described as "rare" in Table 3.1. The major reason for this is that they involve the qualitizing methodology described in Chapter 6: converting quantitative data into narrative data that can then be analyzed qualitatively. While researchers in several of the behavioral and social sciences use the more common quantitizing techniques described in the Type I and IV mixed model studies above, fewer researchers have experience in using the qualitizing methodology to generate profiles and categories.

A recent example of the use of this technique in a confirmatory study is the research of Durland and Teddlie (Durland, 1996; Durland & Teddlie, 1996), which involved the analysis of sociometric data from LSES-IV. Sociometric questionnaires generate data that are inherently mixed: quantitative data in the form of sociomatrices (numerical indices) and qualitative data in the form of sociograms (two-dimensional drawings of relationships among social units).

Sociometric questions typically are very simple; for example, in LSES-IV, the primary question asked faculty members to indicate and rank the top three individuals with whom they had communicated about academic matters in the past two weeks. From this simple question, very complex quantitative sociomatrices (e.g., numerical measures of centrality for the principal, numerical measures of cohesiveness and density for the faculty) as well as qualitative sociograms (pictures of the faculty interaction patterns, including cliques, isolates, and so on) can be generated. In LSES-IV, both sociomatrices and sociograms were used, but it can be asserted that the sociograms produced the most valuable information. The sociograms allowed the confirmation of predictions regarding the interaction patterns that should exist in effective and ineffective schools based on the extant school effectiveness research.

Durland (1996) developed a model, the Centrality-Cohesiveness Model of School Effectiveness, that predicted the types of sociogram patterns she expected from schools varying in terms of a school effectiveness index (SEI; effective or ineffective) and SES of student body (middle or low). These predicted sociograms, or visual representations of faculty interactions, were based on SER theory and research regarding the types of leaders and communication patterns that one might expect to find in such schools. This study used an advanced computer program, KrackPlot (Krackhardt, Lundberg, & O’Rourke, 1993), to generate the sociograms based on the empirical data. The researchers then compared the empirically derived sociograms with the theoretically derived ones that had been predicted in advance.

The sociogram analyses indicated that (a) effective schools were more likely to be dense and centralized ("well webbed"), while ineffective schools were more likely to be spread out and linear ("stringy"); (b) the principal, or surrogate leader, was more likely to be in the center of sociograms of effective schools than ineffective schools, where they tended to be located on the periphery; and (c) there were more isolates and cliques in ineffective as opposed to effective schools.

This component of the LSES-IV study was confirmatory, with hypotheses set up in advance predicting certain types of interaction patterns among faculty members based on the extant SER literature. The data collection and operations were quantitative, with the use of sociometric questionnaires to ask faculty members to rank individuals in their schools in terms of how much they interacted with them. The analyses and inferences for the described component of the study were qualitative, using a computer program (KrackPlot) to generate pictures of the faculty interaction patterns in effective and ineffective schools based on the empirical, numerical data.

Type VI Mixed Design: Exploratory Investigation, Quantitative Data/Operations, Qualitative Analysis and Inference

The LSES-IV study also generated an example of an exploratory study that used the qualitizing technique. As indicated in Table 8.1, LSES-IV was a five-year follow-up study of eight matched pairs of schools. These schools had been initially classified as effective or ineffective based on baseline data from school years 1982-1983 and 1983-1984; their effectiveness status was confirmed through site visits conducted in LSES-III in 1984-1985. While all the schools retained their initial classification status in LSES-III, there was evidence of some naturally occurring school improvement, which was described earlier.

The researchers anticipated that by the time of LSES-IV, some of the 16 schools might be different in terms of their effectiveness status, but they decided not to make predictions in advance regarding these changes. In fact, the researchers were not even sure if the initial dichotomous classification scheme (effective/ineffective) would be adequate to characterize the schools as they had evolved by the time of LSES-IV in 1989-1990.

Instead of predicting the status of the schools in LSES-IV, the researchers decided to use a multidimensional set of quantitative criteria to determine the effectiveness status, and indeed effectiveness categories, for the schools. Teddlie and Stringfield (1993, pp. 84-85) presented data on eight
quantitative criteria for each of the schools: scores on criterion-referenced tests, scores on norm-referenced tests, student attendance, scores on high-inference measures of the quality of classroom teaching, time-on-task in classrooms, stability of faculty, scores on a researcher-completed measure of the overall school climate (hierarchical dimensions of schooling), and changes in the socioeconomic status of the student bodies of the schools.

Based on these quantitative data, the researchers developed a new scheme for categorizing the schools into four qualitatively defined categories: stable more effective, improving, declining, and stable less effective. This component of the LSES-IV study was exploratory, with no hypotheses set up in advance regarding the categories of effectiveness that might emerge. The data collection and operations were quantitative, with the use of eight separate numerical indices of effectiveness. The analyses and inferences were qualitative, using the qualitizing technique whereby numerical data were transformed into narrative descriptions (case studies) of the 16 schools, which then were analyzed qualitatively to produce four emerging categories.

A major purpose of this volume has been to extend the philosophical and methodological “bridges” that are under construction between the QUAL and QUAN research traditions. Both research traditions are very rich and have ardent supporters, some of whom have expended much effort over the past two decades to keep them separated. This effort at reconciliation has been challenging but at the same time rewarding. We had to cover material in several diverse areas, from philosophical foundations to “hands-on examples” of actual research in the field. We hope that we were able to demonstrate that mixing the QUAN and QUAL approaches to research is not only possible but also quite beneficial in many diverse research settings. This point is specifically relevant to graduate students, who are often uncertain about using mixed methods in their dissertation research.

A second point that we have emphasized throughout this volume is the preeminence of the research question over considerations of either method or paradigm. We discussed this early in the volume under the title of “the dictatorship of the question.” The question determines the design of the study, the data collection approach, and so on. The best method is the one that answers the research question(s) most efficiently, and with foremost inference quality (trustworthiness, internal validity). Mixed methods are often more efficient in answering research questions than either the QUAL or the QUAN approach alone. We often hear graduate students say, “If I am going to do my dissertation research using qualitative methods, why should I learn quantitative methods and statistics?” A similar question is asked by students who prefer quantitative methods. We hope this volume has provided some answers for both groups of students through its emphasis on the importance of research questions and on the superiority of mixed methods in reaching comprehensive answers to those questions.

One of the conclusions of this volume is that the paradigm of pragmatism can be employed as the philosophical underpinning for using mixed methods and mixed models, especially with regard to issues of epistemology,