Historically, psychological research has relied heavily on experimental and quasi-experimental studies collecting and analyzing numeric data. When narrative or other unstructured data (e.g., videos, observations, field notes) have been collected, they have been analyzed quantitatively by converting them to numeric indicators. This is still evident in current publications (including this Handbook) and dissertations. Psychological research has also been predominantly hypothesis-oriented, following a process of predicting possible outcomes (relationships, behaviors, etc.) on the basis of literature and/or theories, testing these hypotheses, and making adjustments in theoretical frameworks based on these results. A similar trend is also observable in evaluations of quality in clinical research (for an example, see Tayler & Asmundson, 2008) or other studies using unstructured interviews, observations, and artifacts.

Although a purely positivistic approach is rare to see in portrayals of psychological research, it is helpful to recognize its impacts. The positivist paradigm advocated the presence of objective and discernable realities, leading to the belief that human behavior was predictable from its antecedents, and that distinct causes were identifiable for every behavior (Waszack & Sines, 2003). Despite the appealing nature of such attempts to predict/explain behaviors and relationships, they have been deemed insufficient for almost 50 years. The postpositivist scholars in psychology have provided a much broader approach to research. This approach acknowledges the multiplicity of causes, stresses prediction over causation in research, centers prediction on groups of individuals/events (in contrast to individual-level prediction), acknowledges the subjective role of the participant and the researcher in producing the outcomes, and allows for cultural/value relativity in research. Such paradigmatic changes have been present in psychological research in the form of a “cognitive revolution” (Ross, Lepper, & Ward, 2010, p. 16), with constructs such as cognitive equilibrium/balance/consistency (as in Heider, 1958, and Festinger, 1957, theories/traditions), bias in information processing (Kahneman & Tversky, 1973; Nisbet & Wilson, 1977), and extending to current dominance of social cognition in almost all areas of psychological research.

A consequence of these shifts for psychological research methodology has been the recognition that “targets of any intervention will respond to their interpretations of the program and its consequences—both potential consequences and those that actually take place—rather than those of the designers, implementers, or founders of...”

---

1As an example, see Reis and Gosling’s (2010, p. 100) description of data analysis in “diary method,” highly focused on complex statistical analysis of largely (at least potentially) qualitative narrative data.

2An examination of McKay’s (2008) Handbook of Research Methods in Abnormal and Clinical Psychology would also demonstrate this tendency toward quantification of otherwise qualitative clinical data and observations, as well as standards for assessing quality.

3The prevalence of subjectivism (Ross, Lepper & Ward, 2010) in psychological research during the second part of the 21st century is a clear indication of this shift. Landmark work on “inferential strategies” (Ross et al., 2010, p. 8) led to “greater understanding of the process of ‘construal,’ which is the goal of all work in social perception and social cognition.”
the interventions” (Ross et al., 2010, p. 8, italics in the original). The impact is clearly observable in the way the research outcomes are being assessed and utilized, putting greater weight on the individual-level and contextual impacts on responses, interventions, and sociopolitical issues.

Our own (author’s) methodological growth is another example of such impacts. Examples include conceptualizing the outcomes of mixed methods research as a gestalt (meaning-making in which the investigator connects the inferences from multiple strands of a study to reach a whole that is bigger than the sum of its parts), or applying a social cognition framework to evaluation of quality of mixed methods inferences in our “integrative framework” (to be discussed later in this chapter). Perhaps the strongest impact on us has been in our conceptualization of mixed methods as a “humanistic methodology” that utilizes and extends the “naïve” problem solver’s everyday strategies in systematic and formal ways (Tashakkori & Teddlie, 2010b).

Even with a strong shift toward subjectivism, contextual effects, cognitive processes, and the need for understanding the researcher and the participants’ constructions of events and interventions, the postpositivist paradigm has not been satisfactory to some psychological researchers who have been dissatisfied with its heavy emphasis on hypothetico-deductive process of inquiry (see McGrath & Johnson, 2003). Inductive exploration in the absence of theory and generating explanations from the ground up (grounded theory research) have been the hallmarks of constructivist paradigm in the past few decades. Qualitative research has emerged since the 1970s out of the “methodological tensions” (Eisner, 2003, p. 19) between the old and the new approaches to studying human behavior: “Observing, describing, and understanding became the key terms, and the new business was knowledge building and knowledge generation rather than affirmation or falsification of some previously established hypotheses” (Bamberger, 2003, pp. ix–x).

In the context of presumed conflicts between the two approaches to research in psychology, a note of caution is necessary: Given the fact that relativity of cognitive processes, role of perceptual mechanisms in explaining behaviors, and the impact of culture has a long history in postpositivist psychological research, distinguishing such research from that advocated by constructivists is not easy. “Qualitative inquiry” (Marecek, 2003) has been used for explaining behaviors in laboratory and nonlaboratory studies in psychology for a century. As we discuss later in this chapter, the line between the two approaches to research is at times difficult to find, and perhaps may be best defined as what distinguishes two communities of scholars, rather than two distinct sets of methodologies.4

Reliance on constructivist perspectives, and emergent/emic explanations for behaviors, has not been uniform across subdisciplines of psychology. Clinical psychologists have used unstructured interviews and observations, case studies, and projective techniques to collect data. Although it is still heavily quantitative, cross-cultural psychological research has also used field observations, open interviews, and cultural artifacts as sources of data. Unfortunately, often these narrative/observational data are solely/predominantly statistically analyzed. As an example, the chapter covering recent data analysis advances in social psychology (Judd & Kenny, 2010) in the well-respected Handbook of Social Psychology (Fiske, Gilbert, & Lindzey, 2010) provides evidence of this full reliance on statistical analysis.

Today, the constructivist approach has largely been crystallized methodologically in the form of what everyone calls qualitative research. Although psychology has been slow to change, most social-behavioral sciences have now substantially added qualitative methodology to their training and practice repertoire. That growth has also led to calls for using both qualitative and quantitative approaches/methods (Haverkamp, Morrow, & Pontettro, 2005), in order to avoid the “the overconstraining treatment of complex, dynamic, human systems characteristics of quantitative approaches and the solipsistic epistemological quagmire implicit in the perspectives that characteristically underpin qualitative approaches” (McGrath & Johnson, 2003, p. 46).

Perhaps as a result of these calls, and also due to its multidisciplinary appeal, mixed methods research has steadily gained popularity in the past decade. There is

4There has been a similar conflict over the past four decades between the two approaches across the entirety of the social sciences, starting in more applied fields (e.g., nursing, evaluation, education) and then gradually moving to more basic fields (e.g., psychology, sociology). Arguments have emerged regarding the suitability of mixing qualitative and quantitative methodologies, with some suggesting that these methodologies could not be mixed due to the incompatibility of the paradigms (e.g., postpositivism, constructivism) supposedly underlying them. The result of the incompatibility thesis was the so-called paradigm wars (e.g., Gage, 1989). The compatibility thesis, on the other hand, states that combining different research methods is a good research strategy and denies the contention that these two orientations are “epistemologically incoherent” (Howe, 1988, p. 10).
ample evidence of this in a variety of arenas. Scholars (Bamberger, Rao, & Woolcock, 2010; Plano Clark, 2010) have examined the prevalence of mixed methods writings in at least four interrelated arenas: books, journal articles, dissertations, and funded projects. As of 2011, we know of more than 20 published books about mixed methods. This figure does not include many more books/collections in which mixed methodology appears in at least one chapter (e.g., Denzin & Lincoln’s Handbook of Qualitative Research, 2011). Also, this does not include many introductory research textbooks that now include detailed discussions of mixed methods as a distinct methodological perspective (e.g., Thomas, Nelson & Silverman, 2011).

In 2009, Ivanova and Kawamura (2010) searched for mixed methods journal articles in major indexes. They reported 689 articles published between 2000 and 2008. The number of empirical studies using mixed methods increased from 10 in 2000 to 243 in 2008. The majority of these articles were published in health and medical sciences. Conducting a similar search, Plano Clark (2010) reports that she found more than 700 dissertations self-identified as mixed methods in 2008. She also reports a total of 206 funded projects by National Institutes of Health and 22 other U.S. agencies, with numbers increasing from zero in 1996 to more than 50 in 2008.

It is not possible, at this point, to separate psychological research projects from other behavioral and health sciences in most of these reports (one reason is that many of the journals that publish such studies are not classified as psychology journals). However, one may infer that the rate of utilization of mixed methods has accelerated in psychology as well.5 Strong voices have called for utilizing integrated approaches in psychological research (Dattilio, Edwards, & Fishman, 2010; Fine & Elsbach, 2000; Yoshikawa, Weisner, Kakil, & Way, 2008), advocating “a paradigm shift that would combine qualitative and quantitative approaches” (Torney-Purta, 2009, p. 826).

These voices provide many examples to demonstrate the feasibility/desirability of combining qualitative and quantitative data collection and analysis methods: for example, focus groups and questionnaires (Hahn, 1998); survey

Cases from RCTs constitute a rich source of largely untapped scientific data. Clients in the treatment group are offered a standardized therapy that follows a specific model that has usually undergone considerable development and that is delivered at a quality-controlled, best-practice level. There are comprehensive records of each case: All sessions are audio- or video recorded for fidelity analysis and quantitative measures of important variables (e.g., symptoms and therapeutic alliance) are available pre- and post-treatment and often at follow-up. By the end of a trial, there is a substantial database from which cases of interest can be selected for closer examination and written up as SCSs. Whereas we recommend that some SCSs be routinely published alongside the results of the multivariate analyses, they can in fact be conducted at any point in the future, provided that the recordings and case data are retained with informed consent obtained to allow use of the case material with the usual safeguards for protecting client identity. Of particular interest are differences between treatment responders and poor-outcome clients or nonresponders. (p. 434)

These and other similar proposals are attracting much interest among social-behavioral researchers. An example is a recent article in the Monitor (June 2011) entitled “Mixing It Up,” in which Rebecca Clay reviews the pro- and con arguments about the desirability of case studies in clinical trials.

Why Mixed Methods?

Quantitatively oriented studies are often conducted in larger samples, are predominantly group-oriented (with single-subject studies being the exceptions, because they use small samples and are individual-oriented), and address (direction/magnitude of) relationships between specific...
sets of constructs, rather than conducting “process analysis” (Brady & Collier, 2004). One may also infer that these studies are more “problem-oriented” (Wilson, Aronson, & Carlsmith, 2010) than “process-oriented.” In comparison, qualitatively oriented studies are often conducted in small samples, are individual-oriented, and are predominantly concerned with process and the context of behaviors. The challenge, and the rewards, of conducting mixed methods research is the opportunity to use both approaches, potentially leading to far greater understanding of the phenomena or behaviors under investigation.

As a demonstration of the benefits of combining what she calls person-centered and variable-centered research approaches to studying adolescents, Torney-Purta (2009) reviewed a large-scale multinational study of youth attitudes. The first strand of the study consisted of the usual in-depth analysis of youth attitude surveys in 10 countries. The second (person-centered) analysis consisted of cluster analysis—identifying groups of youth in various countries based on their political tendencies (e.g., conventional, alienated, indifferent)—followed by constructing a profile of each group in each country. These profiles were then compared with each other in the context of countries’ political/cultural differences (e.g., comparing Eastern European countries with Western European, or the United States with European countries). The comparisons provided insights that would have otherwise been missed if the quantitative survey outcomes were the main source for drawing inferences.

Torney-Purta makes further observations regarding the benefits of combining the two approaches:

An advantage of research that combines qualitative and quantitative measures is that audiences outside of the academic world can often be more readily convinced of a need for action if they can see research findings as part of a context rather than as an abstract statistical procedure. Calling attention to groups of individuals having particular profiles within and across countries can aid researchers in interpreting the information gained from cross-national summary statistics. When it is possible to see a cluster of individuals adolescents who remind them of young people they know, those involved in policy and practice are much more likely to take action than when they are told only about averages and statistical trends. (p. 834)

Although she does not present a formal approach to qualitative analysis, we think her strong call for incorporating the cultural context in interpreting psychological research findings fits well within the framework of mixed methods.

We would like to point out that although post-experimental debriefing (Wilson et al., 2010) has been advocated and extensively used in well-known experimental studies, these studies are not necessarily mixed. Debriefing has been used to inform the participants about the presence of any deceptions, identify the degree of awareness of experimental expectancies/hypotheses, and assess the potency of interventions. This is not the same as conducting in-depth interviews (paired with possible direct observations and field notes) and subjecting the data to formal content analysis to identify the processes underlying the impacts of interventions (or lack thereof), but debriefings do point out the importance that qualitative methods could have in psychology, especially with regard to better understanding how participants experienced their experimental settings.

This is also applicable to nonexperimental studies. Including one or more open-ended questions in a questionnaire does not make a survey study mixed, if the responses to open-ended items are converted to numeric indicators and analyzed statistically. On the other hand, this would be a mixed methods survey study if the narrative data from these open-ended items (or interviews) are content-analyzed to identify themes that might otherwise be missed by closed-ended items, to identify possible cognitive processes underlying the responses, to discover divergent points of view, or to enable the investigator to interpret the quantitative findings in the context of personal interpretations. An example of such a scenario, using interviews, may be found in Yoshikawa et al.’s (2008, p. 346) description of inconsistent findings:

Beliefs, goals, and practices are particularly interesting when they are not congruent. The combination of quantitative and qualitative evidence can shed light on why this is so. In a recent study, Hughes et al. (in press) examined beliefs regarding the importance of various ethnic and racial socialization practices, as well as frequencies of those practices themselves, in a sample of 210 Chinese, African-American, European-American, and Latino adolescent–parent pairs. Both survey and semi-structured interview data were collected from both teens and parents. The researchers first uncovered discrepancies in their survey data between levels of beliefs and practices within participants as well as levels of beliefs or reported practices across the teen and parent in a particular family. The semi-structured interview data helped shed light on why the discrepancies occurred. For example, it appeared that routine, everyday activities (revolving around food, books, films, or artifacts, for example) were often identified as associated with ethnicity but not perceived as examples of intentional cultural or ethnic socialization.

Utilizing Mixed Methods in Psychological Research 431
Similar advantages of detecting discrepancies have been reported by others. Davidson, Wieland, Flanagan, and Sells (2008) have provided a number of such studies from research on psychopathology (for example, Davidson & Stayner, 1997; Miller, 1994). In these studies, qualitative interviews of individuals who had previously been diagnosed with borderline personality disorder or schizophrenia revealed a very different self and/or social identity than otherwise measured by more structured clinical interviews: “In contrast to not knowing who they were, they...were able to indicate instances when they could be themselves and instances they could not” (Davidson et al., 2008, p. 261).

In a call for integrated studies of infant sleep, Middlemiss (in progress) demonstrates how our understanding of mother-infant sleep arrangements changes when the traditional etic-oriented quantitative research methods are combined with more emic-oriented approaches involving qualitative cultural observations and in-depth interviews with parents. In her study of New Zealand mother-infant sleep patterns, quantitative results indicated a positive correlation between mothers’ physiological stress at infant’s transition to sleep, self-reported stress, and mothers’ report of infants having trouble learning to sleep during the night. Contrary to this accepted pattern, qualitative data (ethnographic interviews in their homes) indicated that mothers who rated themselves as comfortable with the idea of shared sleep expressed more stress about nightwakings than others. However, these mothers attributed source of this stress not to infants’ sleep behavior but to their fatigue from daytime care activities. Without an integration of the emic/qualitative data with quantitative results, she suggests, our understanding of shared infant sleep would be incomplete, or misguided.

**Mixed Methods Research in Psychology: A Historical Overview**

In 2003, Cindy Waszack and Marylyn Sines reviewed and classified the studies in psychology that could be identified as mixed. They grouped these studies in three broad categories of sequential, parallel, and complex mixed methods. They observed that the majority of sequential mixed studies started with a qualitative phase that was followed by a quantitative one. Some of the work reviewed by them actually consisted of a program of studies conducted over time, leading to a mixed sequence of research projects that were focused on the same question or theoretical framework. Others were single studies with multiple phases/strands.

In their sequential design category, they examined eight broad sets of studies, six starting with a qualitative phase, two with quantitative:

- Qualitative strand followed by quantitative strand:
  - Laboratory tests of theories developed from personal experiences.
  - Laboratory test of theory originating from initial interviews.
  - Cognitive dissonance and the end of the world (as we know it).
  - Psychological anthropology and the Oedipus complex.
  - Gender differences in sexual motivation and strategies in Cameroon.
  - Grounded theory research on cognitive development and construction of maternal role.

- Quantitative strand followed by qualitative strand:
  - Gender norms among adolescents in Jamaica.
  - Adolescent dating aggression in the United Kingdom.

Table 15.1 summarizes some of the sequential studies reviewed by them.

Waszack and Sines’ review of parallel mixed studies also included multiple study projects that spanned over a long time, as compared to multiple strands within one research project. Broad examples of these were McGuire’s spontaneous self-concept studies, Milgram’s obedience studies, and Sherif’s Robber’s Cave Experiment. Table 15.2 summarizes these.

Finally, Waszack and Sines identified some studies as “complex mixed methods,” including combinations of the two other simpler types of designs. Zimbardo’s (1969) prison and de-individuation study, and Tolman and Szalach’s (1999) dimensions of desire were summarized as examples of this type of studies (see Table 15.3).

Although these examples demonstrate the historical prevalence of groundbreaking mixed studies in psychological research, some of the projects reviewed were “programs of research” with many iterations across decades of work. This is certainly consistent with some scholars’ definition of mixed method research (e.g., Morse & Niehus, 2010). As we discuss later, many of the classic and current...
Cialdini developed a lab test of two competing theories that initially originated from his own personal experiences with the United Way. Was able to confirm one and discard the other.

Freedman and Fraser, 1966

Lab experiments tested the theory that complying with a small request will lead to compliance to a larger one. The theory was originally developed via in-depth interviews with former U.S. POWs who experienced brainwashing in Korea.

Festinger, Reicken, and Schacter, 1956

A small group who believed in an end-of-the-world prophecy generated qualitative data that formed a program of experimental studies on cognitive dissonance. When the prophecy failed, Festinger et al. tested several hypotheses about the impacts of belief disconfirmation.

Johnson and Price-Williams, 1996

A qualitative study of classic myths and folktales from around the globe that identified different types of roles a father takes in these stories. Multiple dimensions were then coded in 164 folktales, and were then linked in quantitative analysis.

Calves, Cornwell, and Envegue, 1996

Four focus group discussions explored the sexual experience, strategies, and motivations of males and females.

Flanagan et al., 1995

In a study of the relationship between adolescent cognitive development and the concept of maternal roles among young mothers, data from group and individual interviews with 42 teenage mothers were content-analyzed using the constant comparative method. Five hypotheses were formed on the basis of emerging categories, and were statistically tested after the data from subsequent interviews with 25 adolescent mothers were coded.

Eggleston, Jackson, and Hardee, 1999

A survey of 945 Jamaican youth demonstrated discrepancy between boys’ and girls’ sexual experiences. Eight focus groups were conducted, each consisting of eight same-sex/age members, asking them about their understanding of sexual issues. The combined findings made it possible to construct more complete and meaningful picture of students’ beliefs and behaviors.

Hird (2000)

Questionnaires followed by focus group discussions and interviews provided different types of results about aggression that is experienced by girls and boys in heterosexual dating situations. Qualitative findings found less support for symmetry of aggression.

mixed methods studies deviate from that trend in that multiple types of data are collected and/or analyzed in a single project. In historical context, Jahoda, Lazarsfeld, and Zei
tel’s (1971) report of the 1930 “Marienthal” project is a classic example in which using diverse methods of study and data collection/analysis techniques provided a rich profile of a chronically unemployed community in Austria. In the original study, published first in 1933, the investigators skillfully utilized qualitative and quantitative methods and data, blended in an effort to reach a comprehensive understanding of a case. The 1971 report provides a translation of the original, and also includes one of the earliest calls to develop systematic procedures for incorporating the two approaches. Marienthal was a small village in Austria during the Great Depression.

In the introductory chapter of the English translation, Jahoda et al. clearly put forth their reasons for using such an eclectic approach/design: “Our idea was to find...
procedures which would combine the use of numerical data with immersion (sich einleben) into the situation. To this end it was necessary to gain such close contact with the population of Marienthal that we could learn the smallest details of their daily life” (pp. 1–2). The outcomes of such a diverse scope and eclectic approach was a rich and comprehensive dataset that included multiple layers of the social system and multiple perspectives to the issues of interest to the study, as demonstrated by the following list:

- Family files.
- Life histories.
- Time sheets.
- Reports and complaints made over the past few years to the industrial commission of the district.
- School essays written by primary and secondary school children.
- A prize essay competition by adolescents: “How I see my future.”
- Meal records families kept for one week.
- Christmas presents received by 80 small children.
- Conversation topics and activities in public bars.
- Parents’ problems in bringing up their children (notes taken in the doctor’s consulting room).
- Information from teachers on the performance of their students.
- Money spent at the tavern and a variety of other places.
- Archival data, including household statistics (Jahoda et al., 1971, pp. 4–5).

Diversity of data, data sources, and data collection techniques, matched with exceptional analytic and inferential skills of the investigators, provided a detailed and rich profile of the setting and the people under study.

Planning Mixed Methods Psychological Studies

Planning a mixed methods study is often in the context of at least two relatively distinct scenarios. In one, the study is planned with multiple strands ahead of time in order to answer questions that clearly need a mixed methods design. In the other, the need for a mixed methods design emerges sometime during the course of the project, such as when the results are unexpected or not as meaningful as one expected them to be. The difference between the two contexts is reflected in the research purpose and question(s), and must be deliberated upon and reported clearly. In both, the research question(s) clearly demonstrate why it is necessary (or preferable) to use both qualitative and quantitative data/approaches. In both, one must decide which general type or family of mixed designs would make it possible to make credible inferences to answer the research question(s).

Also anchored in research purposes and questions, the researcher must make other types of decisions about an optimal mixed design. One is about the necessity of collecting new data in a subsequent strand, as compared to converting and analyzing the data that were collected in an initial qualitative or quantitative one. The other is a decision regarding the type of data sources (sample) to use for different strands of the study. As we discuss next, choices might include using the same sample of observations, selecting a subsample or, alternatively, collecting data in a different sample or a different level of a social organization (in a hierarchical sample).

Mixed Methods Designs

Like every other scholarly field in transition, the field of mixed methodology has undergone tremendous transformation in the past three decades, certainly since the late 1990s when major books and manuals were published (Bamberger, 2000; Creswell, 1994; Greene & Caracelli, 1997; Newman & Benz, 1999; Tashakkori & Teddlie, 1998; among others). Mixed methods designs have advanced into complex and systematic ways of doing research. They are also evolved in the form of multiple typologies that are not necessarily consistent with each other. Our own conceptualization of mixed methods designs has also evolved since our early work in this area (Tashakkori & Teddlie). Teddlie and Tashakkori (2009) have conceptualized mixed methods designs in four broad families of sequential, parallel, conversion, and fully integrated. The fourth family (fully integrated) includes more complex designs that employ a combination of the other three. This classification simultaneously takes the process, purpose, and expected outcomes of the research project into consideration. Within each design family, there is variation on the basis of units of study (individuals, social units, or both). The fourth family (fully integrated) also includes multiple strand designs utilizing more than one sampling framework.

Sequential Mixed Methods Designs

In sequential mixed studies, the questions, sample selection criteria, data, or other components of a strand are based on the findings of the previous strand. Obviously, the second strand cannot be initiated until the data of the first strand are analyzed and the results evaluated. For example,
until the investigator determines which participants in a randomized clinical trial have not responded as expected, it is not possible to follow up. In the postexperimental case comparison method (Dattilio et al., 2010), investigators compare good outcome and bad outcome cases, in order to understand the process through which the experimental treatment has failed and/or the alternative treatment has led to effects expected of experimental participants.

The questions of a second strand might also emerge from the results of the previous one, even though the investigator did not initially intend to conduct a mixed study. Wesely’s (2010) study of learning motivation, to be discussed later in this chapter (data analysis), provides an example. Figure 15.1 depicts this type of design. Unexpected findings in an experiment or a survey study might necessitate such a follow-up. For example, based on the results of the first strand of a study that tries to identify successful and unsuccessful modes of coping among a group of individuals, the investigator might decide to select, for detailed case studies, several individuals who are coping successfully with daily stressors and several who are not.

An example of this type of design can be found in Watson, Goldman, and Greenberg’s (2011a, 2011b) case-comparisons of a successful and an unsuccessful emotion-focused therapy for depression in a randomized clinical trial consisting of 16 to 20 sessions (as other examples, also see the case studies by Buckrell and McCain, as well as commentaries about this type of study, in the same issue of Pragmatic Case Studies in Psychotherapy). A wide variety of data were collected in a sample of 38 depressed clients, including their self-report responses to instruments such as Symptom Checklist, Inventory of Personal Problems, and Rosenberg Self-Esteem Inventory. The data also included the therapists’ notes and responses to formal assessments (e.g., therapist postsession questionnaire; posttherapy assessment interview).

The respondents had been recruited by an ad in the local newspaper, seeking participants for a study of treatment for depression. Each participant had been randomly assigned to one of two different methods of therapy. The two cases presented in the 2011 (a and b) reports were both from the emotion-focused therapy group. One (Eloise) was selected because she was a model of success for the therapy technique, and the other (Tom) because he was a model of failure. Tom’s scores on the Beck Depression Inventory (BDI) dropped modestly, from 24 to 14 after 16 sessions of (randomly assigned emotion-focused) therapy. The authors (p. 272) note that a score of 15 or higher on the BDI indicates severe depression. A score of 14 was in the upper limit of mild depression classification. Eloise’s BDI scores, on the other hand, were initially higher (40), and dropped to 0 well before the 11th session of therapy. The two also showed differences in other indicators of adjustment.

Comparison of Tom and Eloise’s development history, social support networks, strategies for dealing with emotions, and other adjustment-related variables revealed two different patterns of adjustment. Eloise became rapidly engaged in the therapeutic process, effectively utilizing the type of strategies that the emotion-focused therapy aimed to strengthen (“It’s like opening the windows and letting the fresh air come in”). Tom (“trapped in the tunnel”), on the other hand, started therapy with skepticism, had difficulty experiencing and expressing his emotions, and was not successful in establishing a social support network. Detailed comparison of the two cases confirmed the predictions based on current theories/models of psychotherapy; that is, the differences were highly consistent with the factors theoretically expected to facilitate or hinder the therapeutic change process. Adding a case-comparison strand to the randomized control trial provided insights about the factors, processes, and the contextual variables

---

**Figure 15.1** Sequential mixed methods design in which the research questions of the second strand emerge from the findings of the first (QL = qualitative, QN = quantitative, order may be reversed)
that impacted the success (or lack thereof) of the intervention. Adhering to a traditional one-strand experimental methodology would have deprived the investigators of such insights.

Another example of emergent questions is the ethnographic follow-up study of a subsample of the New Hope Project in Milwaukee (Duncan, Huston, & Weisner, 2006). The project, originally started in 1994, included a sample of more than 1,300 adults who applied for participation in the project, and met a few basic conditions such as eligibility for welfare programs, providing proof of at least 30 hours of work per week, and residence in the poorest area of the city. They were randomly assigned (selected) to experimental (New Hope Project, n = 678) and control (governmental welfare recipients, n = 679) groups. Both groups of adults continued their eligibility for all other federal, state, and local programs. The experimental group received incentives such as income supplement insurance, child care, and any other needed assistance (e.g., job counseling) that facilitated earning more than welfare programs in the 3 years of participation. Those with children received a more generous package of assistance, putting them substantially above the poverty level.

Evaluation of the effectiveness and impacts of the program started in 1996 (after 2 years of participation) by collecting data regarding family transitions, job histories, and economic status of the adults in both groups. More than 800 participants in the two groups had young children between the ages of 1 and 12. In a subsequent investigation (the Child and Family Study), the effects of the program on family life and children’s development were evaluated. The data included children’s school performance, psychological well-being, and behavior problems as reported by teachers, parents, and the children themselves. The investigators also collected data on parents’ self-reported stress, hopefulness, depression, and quality of their relationship with their children.

Although the results of the evaluation (and subsequent analyses since then) were generally favorable to the experimental program, many inconsistencies and puzzles were also reported about the impacts. In order to further understand the effects of the program, and also to clarify the factors that facilitated or hindered the program effectiveness, a follow-up qualitative (ethnographic) study was conducted in a sample of 44 families, randomly selected from the more than 800 participants who had young children:

\[T\]he ethnographic evidence suggests that the flexibility of New Hope benefits allowed families to use them in many ways that met particular family needs. Hence, there is no one pathway by which effects occurred. Some parents had consistently higher incomes because of New Hope; others had new social contacts from going to work; and others experienced improved psychological well-being and pride in their ability to support their families.

The survey and the family stories from our intensive interviews helped to solve some of the puzzles and to identify who was helped, who was not, and why. The New Hope offer made a big difference for some people but it was not a good fit for everyone. Some parents refused to entrust their children to the care of someone other than a family member. Many parents worked evenings and weekends, when few child-care centers or licensed home settings were available. The child-care subsidy was therefore of little use to them. Others had personal difficulties that kept them from regular work. The vagaries of low-wage jobs also made it difficult for some people to use the New Hope benefits. Some had irregular and unpredictable work schedules that did not allow them to meet the thirty-hour weekly work requirement (Miller, Huston, Duncan, McLoyd, & Weisner, 2008, p. 13).

Another example is a study of post-Katrina mental health services in Louisiana (St. Bernard Parish) by Mitchel et al. (2008), constructing a mental health questionnaire on the basis of the clinical observations and interviews by the first two authors more than a year earlier. Mitchel was the “sole mental health professional” (p. 67) during the two weeks of her deployment in St. Bernard Parish in December 2005, providing services in a makeshift clinic. The first set of data consisted of mental health intake interviews, progress notes, treatment plans, anecdotal observations, and field notes. There was ample evidence indicating the widespread prevalence of symptoms of grief, depression, posttraumatic stress, adjustment disorders, and psychosis. Data were recorded for 2,150 patients who were seen in the medical clinic during 22 consecutive days, 4 months after Katrina. Among that sample, 212 were identified as needing mental health assistance, showing symptoms of depression, suicidal ideation, and other problems. Investigators’ experiences in the field and examination of the records pointed to numerous ambiguities and inconsistencies in the behaviors and coping mechanisms in the first few months after Katrina. Among them were inconsistent individual reactions to the disaster, institutional chaos, and the need for comprehensive knowledge about prognosis and predictability of the symptoms by the providers. Eight specific questions were asked, and answers were sought in the next phase of the study, in 2006.

Returning to St. Bernard Parish more than a year and a half later, the investigators collected data to better understand the impact of the large-scale disaster on the mental health of the victims. In addition to the archival data
collected by clinics and agencies, the authors developed a questionnaire measuring the participants’ self-reports of experiences in four time periods since Katrina, current problems, and other variables. Data (responses to close-ended and open-ended items) were collected from individuals 18 years or older who were being treated for routine (nonurgent) medical (nonmental health) problems. The data were statistically analyzed and reported descriptively, along with a full list of symptoms, problems, and typical responses to them.

The findings were interpreted in the context of qualitative results of the first strand (18 months earlier). Among the emergent conclusions was the finding that post-Katrina symptoms were 10 times more prevalent than expected. Almost two thirds of the respondents had developed new symptoms, and the majority of the individuals were still experiencing mental health difficulties. Furthermore, many symptoms recurred long after the disaster struck. Qualitative results pointed to participants’ lack of hope (“no obvious ‘light’ at the end of the tunnel,” p. 75). The results also indicated that a majority of individuals in the follow-up sample needed mental health therapy, but only 20% had received any.

As discussed earlier, in a sequential study, it is also possible to plan for a second strand ahead of time, knowing that a certain outcome would be necessary before the second phase of the study can be implemented. For example, in order to select a sample for the second (or third) strand of a study, it is necessary to collect and analyze data, and use the outcome to select a sample of individuals for the next strand. As an example, it is possible to predict that some members of the experimental group will not respond to an intervention, although it is not possible to predict who exactly will be in that category. The questions of the second strand are known ahead of time: Why do some members of the experimental group not respond to treatment as predicted? These research questions might be modified as the study progresses. An in-depth qualitative inquiry provides answers to this question. Figure 15.2 demonstrates this type of study.

An example of exploring such preplanned questions in subsequent strands of a research project is the Hitchcock et al. (2006) study involving the development of “ethnographically informed instruments” (p. 19) to measure adolescent self-perceptions in Sri Lanka. Interview questions were developed by a group consisting of the authors, an educational sociologist and teacher educator, and a child psychiatrist. Questions included types of academic problems, family issues, and related adjustment difficulties. Interviews were conducted in 18 schools, with 33 groups of students, 18 groups of teachers, and the 18 principals. Qualitative data were content analyzed, with specific attention to respondents’ expressed expectations of male and female children. Three broad categories of themes emerged in the analysis: suitable behaviors, unsuitable behaviors, and personal/interpersonal needs. Broad gender differences in expected competencies were found within these categories.

Qualitative data also provided a list of culture-specific constructs pertaining to self-perceptions, mental health issues, gender-specific concerns, and expected competencies. The culturally specific competencies list was used to construct a Likert-type scale, with 3-point response options ([have] a lot = 3, some = 2, not at all = 1).

Quantitative data consisted of 611 adolescents’ responses to this constructed culture-specific competency scale. The respondents were students in grades 7 to 12 of six schools, selected purposefully to represent the range of socioeconomic, ethnicity, and religious background of the population. Quantitative analysis consisted of factor analysis of responses to the questionnaire. Multivariate group comparisons were made to test gender differences in the factor scores, followed by comparisons on select items.

![Figure 15.2](image-url)
Hitchcock et al. present and discuss the results of two strands separately, then link them to each other. For example, gender differences in the factor structures and factor loadings were compared with the pattern that emerged in the qualitative analysis. Also, qualitative findings were used to explain unexpected and/or inconsistent findings of the statistical analysis.

Preplanned sequential mixed design was also used in Wall, Devine-Wright, and Mill’s (2008) study of the role of perceived behavioral control in shaping individuals’ normative beliefs about modes of transportation. In the first strand of the study, a sample of more than 1,000 students and university staff responded to a questionnaire regarding their use of cars and other modes of transportation, as well as items corresponding to the two theories of Planned Behavior and Norm-Activation. Among the respondents, 539 reported driving their cars to campus at least 80% of the time. Quantitative analysis of their responses included factor analysis of the survey items identifying five oblique factors: awareness of consequences, attitudes, perceived behavioral control, subjective norms, and personal normative beliefs, explaining 58% of the variance of the items. Analysis also included separate logistic regression, predicting intentions from other components within each of the two theories.

The qualitative strand was planned ahead of time, with interview as the data collection method, including questions pertaining to the component of the two theories under consideration (Planned Behavior; Norm-Activation), with the goal of finding shared positions (those articulated by more than one person) regarding commuting modes. Among the 79 participants of the first strand who agreed to continue for the second strand, 24 were selected on the basis of standardized regression residuals. This subsample represented four groups of individuals, each demonstrating the best fit of their data to each of the two theoretical models, as represented by regression residuals (i.e., fit with Theory of Planned Behavior, fit with Norm-Activation theory, both, neither).

The 24 participants were interviewed, using a semi-structured format consisting of 12 questions that varied in focus (from general to specific). Using Template Analysis (King, 2011), they content-analyzed the responses. Analysis focused specifically on identification of shared themes (those that identify themes strongly expected to be relevant to the analysis), and continue with emergent ones.

The authors discovered three ways in which the personal normative for reducing car use and perceived personal control over mode of transportation were invoked together. Qualitative findings suggested that the concepts of both theories were salient for individuals, and the variables that they perceived shape their behaviors were consistent with the statistical models. Consistency between the findings of the two strands, despite methods and perspective, provided stronger inferences about the impact of personal normative motives on behavior. The von der Lippe (2010) study of young men’s transition to fatherhood demonstrates another sequential design with emergent questions. Data came from the participants in the Rostock Longitudinal Study, which started with all newborns in the Rostock obstetrics ward, and continued with a follow-up of 294 children in 1972. In 2002 to 2003, 206 of the original sample (now at age 32) were followed up, identifying 117 males who were first-time parents. The survey (quantitative) data came from this sample.

From the sample of childless men, 80 whose current place of residence was known were invited by mail to participate in interviews (offering each $25). Among them 25 responded (31%), and 20 were interviewed. Qualitative strand of the study used in-depth interviews of these 20 childless men, each lasting between 70 and 150 minutes.

Quantitative data were analyzed via hazard regression (survival analysis), “to model the transition into parenthood of childless individuals over time” (p. 206). Among the findings that were difficult to interpret, it was discovered that men who value intimate relationships showed a reduced propensity to fatherhood. The qualitative data were content analyzed by identifying themes represented by every sentence and paragraph of the interviews, leading to several dozen codes/themes. These codes were then grouped (co-coded) in terms of degree of representation of several psychological attributes (motives, interests, goals, self-concept, etc.). The final outcome of these analyses was
the identification of two core categories explaining men’s desire for having children: *Developmental Perspective of the Self* and *Evaluations of Social Objects*.

The findings of the two strands were integrated at the end (“juxtaposition,” p. 210) by asking “Which quantitative determinants of childbearing also play a role in men’s subjective motivation for it?” (p. 210). Conclusions were drawn by subjecting “every result from one approach to the scrutiny of the respective other approach” (p. 217). For example, they (p. 214) concluded that “men who report a particularly high evaluation of the quality of their relationships with friends” (from the qualitative analysis) had a “reduced risk [probability] of parenthood” in hazard regression models.

As demonstrated by examples throughout this chapter, sequential mixed studies might collect their data in a variety of ways in one or more samples throughout the entire study:

- A different sample of individuals might be selected for each strand. The Mitchel et al. (2008) study described above is an example.
- One strand might use a subset of individuals from the previous one. The Wesely (2010) study of learning motivation, presented below, is an example. As another (hypothetical) example, within a large school district, initial survey results might be used to identify teachers with extreme (high or low) predicted burnout probability (e.g., regression residuals). The investigator conducts in-depth interviews in a subsequent qualitative strand.
- The strands might use data from two or more levels of a social system or organization, for example, children in one strand, parents in another. An example might be an extension of the above, if the schools with highest or lowest teacher burnout predicted scores are selected, and in-depth interviews are conducted with the principals. Hitchcock et al. study (2006), described in detail above, is another example in which group interviews were conducted in a multilevel sample consisting of principals in 18 schools, a group of teachers in each school, and 33 groups of students.

Regardless of sampling strategy, in all sequential mixed designs, the sample, procedures, instruments for data collection, or even data analysis techniques of the second strand are dependent on the results obtained from the first.

**Parallel Mixed Methods Designs**

In parallel mixed studies, the investigators(s) collect and analyze data in two relatively independent strands (phases) that are planned and implemented in order to answer a research question requiring both qualitative and quantitative approaches. At least one strand involves quantitative-oriented questions, data, and/or analysis, while the other uses a qualitative approach. The findings and inferences emerging from these two strands of the study are integrated at the end, as a higher-level answer to the initial multifaceted research questions. Figure 15.3 is a graphic representation of this type of study.

Edwards and Lopez’s (2006) study of perceived family support, acculturation, and life satisfaction among Mexican-American youth is an example of a parallel study using data from the same sample. In this survey study, a Likert-type questionnaire provided the quantitative data, and an open-ended item asking the participants to write about the perceived factors contributing to their life satisfaction provided the qualitative data. Each type of data was analyzed accordingly, and the findings integrated. The investigators found the family to be the most important perceived contributor to the happiness and cultural adjustment of Mexican-American (but not white non-Hispanic) adolescents.

Another example is the Castro and Coe’s (2007) study of acculturation and alcohol use among young women.
who lived close to the Mexican-American border. Their sample consisted of 77 women who were selected from 318 participants in a larger study of perinatal care services in Arizona. Qualitative data were collected via face-to-face interviews, responding to open-ended questions. Quantitative data came from the same sample’s responses to a questionnaire on beliefs and acculturation. A one-hour structured interview, consisting of 14 sections was conducted with each participant. Interview protocol included items measuring women’s background, level of acculturation, pregnancy-related beliefs and behaviors, cultural involvement, and other formal scales measuring family traditionalism and lifestyle.

Qualitative data were content-analyzed via the Text-Smart (SPSS) program leading to emergent themes in a four-step strategy: (1) excluding irrelevant terms (those with nonsubstantive meaning); (2) creating aliases (words that share similar meaning, for example, hot, sizzling, and scalding all mean very hot); (3) generating categories based on frequency of response, co-occurrence, or both; (4) iterative analysis to generate complex categories, consisting of two or more words/concepts that were connected to each other in the narrative (for example, a traditional person listens to her elders). A final (fifth) step consisted of axial (dimensional) coding by assigning codes based on intensity of occurrence within each of the categories of texts (0 = not present; 1 = tangential reference; 2 = clear solid reference; 3 = emphatic reference). The rating was done independently by three undergraduate students. The team constructed final codes after reaching consensus, if there were differences between the raters. These quantitized codes were then used in logistic regression analyses, separately predicting pregnancy self-care beliefs alcohol abstinence from other variables. The investigators analyzed the correlation between the results of the quantitative scales and the quantitized data representing the qualitative themes. Qualitative results were also used to interpret the unexpected or contradictory findings of the quantitative analyses (such as negative correlations where positive correlations were expected, and the predictive failure of the scaled cultural variables such as acculturation and family traditionalism). The authors made conclusions on the basis of each type of result, and integrated these conclusions to form a comprehensive understanding of the relationship between acculturation, alcohol use, and perceived family support. As an example:

[T]he thematic variable of machismo was not significantly associated with either of the outcome variables. It was nonetheless noteworthy that many of our participants commented on the detrimental influences of negative machismo on the family, criticizing the abusive behaviors of macho men, as well as the submissiveness of very traditional women. (p. 282)

Townsend, Flores, and Findling’s (2010) study is an example of parallel mixed methods study in which a subsample was selected for the second strand. The data of the first strand came from 122 adolescents who met the DSM-IV Axis I diagnosis, as well as certain other family and personal attributes. They responded to the Drug Attitudes Inventory (DAI) measuring attitudes toward psychotropic medications. Concurrently, in-depth interviews were conducted with 20 of the adolescents as well, asking about their subjective experiences with medications. The survey data were factor analyzed (both exploratory and confirmatory). Qualitative data were content analyzed (line-by-line codes, as well as constant comparative), identifying different themes representing youth experience with medication. The outcomes of these analyses were integrated by comparing and contrasting the obtained dimensions, using qualitative findings to interpret the (relative) lack of adequacy of the seven-factor solution from the quantitative factor analysis. Among the final conclusions and recommendations was the necessity of rewording items to make them more relevant to adolescent experiences.

Conversion Mixed Methods Designs

Many clinical studies in psychology code the data that are collected via unstructured observations and interviews, and analyze them quantitatively. This is also the main attribute of conversion mixed methods designs. However, in conversion mixed methods designs, the data are analyzed both quantitatively and qualitatively, and inferences are made on the basis of both sets of results. For example, in Castro and Coe’s (2007) study discussed earlier, codes were assigned to concepts in terms of their intensity of concurrence, and were analyzed quantitatively (correlations, regression). The themes emerging from content analysis of narrative data were also compared and contrasted or linked to the results of quantitative analysis of these coded variables and the closed-ended measurement scales. In other words, a simple conversion of qualitative data for statistical analysis is not sufficient for identifying the study as a conversion mixed study. Although a QUAL-QUAN conversion is more prevalent in psychological research, QUAN-QUAL conversions are also feasible. For example, the narrative profiles constructed on the basis of MMPI scores of a group of individuals might be content analyzed to see recurring themes.
In conversion studies, the data collected from the entire sample may be converted and analyzed separately (as an example, see Castro and Coe’s 2007 study discussed earlier), or, a second analysis might be based on a subset of the first strand. Obviously, the conclusions gleaned from both sets of findings are compared, contrasted, and integrated for better understanding of the phenomenon under study. Figure 15.4 presents a general image of this family of designs.

An example may be found in Devine, Reed-Knight, Loiselle, Fenton, and Blount’s (2010) study of posttraumatic growth in young adults who had experienced childhood illnesses. A sample of 62 young adults responded to advertisements and the university research pool. Sixty of these participants completed a questionnaire including demographic characteristics and a screening item asking them if they had experienced any serious and/or chronic illnesses during their childhood. They also responded to two other quantitative scales measuring their experiences with illnesses and their impacts (Posttraumatic Growth Inventory; Impact of Event Scale-Revised). Responses were analyzed, using a variety of techniques ranging from descriptive to correlation and analysis of variance. Pre-planned stepwise regression analysis was used to predict posttraumatic growth. Accounting for 47% of the variance of the posttraumatic growth, the predictors were perceived severity, illness status, and posttraumatic stress symptoms.

Qualitative data consisted of responses to three separate open-ended questions by 59 of the original 62 participants. The first question was the most general (“Tell us about your experience with a serious childhood illness”), followed by two more specific questions (“Describe any positive aspects of having a serious childhood illness”; “Describe any negative aspects of having a serious childhood illness”). Responses (narratives typed by the respondents) were content-analyzed in a two-stage process: First, identify 44 general categories of experiences (themes, codes), then reduce them to 15 (6 positive and 9 negative outcomes experienced). Two raters rated the narrative for intensity in each of these 15 categories, thereby quantitizing the narrative responses (0 = not present; 1 = mild, lacking description; 2 = moderate/severe, significant description). In case of inconsistency between the two raters, a third rater reconciled them. Content analysis of the results revealed salient positive and negative experiences in relation to having had an illness in childhood. Positive shifts in perspective and frequent medical requirements were among the emergent themes. In addition to narrative summaries of the categories/themes, the authors also presented descriptive statistics for these coded indicators, followed by linking the qualitative and quantitative results:

"The quantitative analyses showed that disease factors and ongoing distress “set the stage” for growth but did not reveal the processes associated with the development of growth following trauma. The qualitative data provided further, more detailed, descriptions of factors associated with growth that were most salient to young adults as a result of experiencing a serious illness. When asked to describe their experiences with a serious illness . . . participants frequently wrote about disease factors, including explanations of the physical/health problems, frequent medical requirements, the severity of the condition, restrictions in daily activities due to the illness and/or its treatment, and negative emotional responses to the experience. (p. 346)

Fully Integrated Mixed Methods Designs

This type of design utilizes a combination of the three basic types discussed above (parallel, sequential, conversion), in an interactive and dynamic manner. For example, in a parallel mixed methods study the qualitative data might also
be quantitized for another round of analysis (thus resulting in two sets of statistical analysis and one set of qualitative). Alternatively, in a sequential design, two parallel studies (one qualitative, one quantitative) might be conducted at Time 2, in order to further expand/understand/explain the findings of the first. The parallel strands might use subsamples of the previous strand, or select different samples as appropriate for answering the sequence of research questions. See Figure 15.5.

The Barg et al. (2008) study of loneliness and depression among older adults (65-plus years old) is an example of this type of design in which integration occurred iteratively in sampling, data collection, data analysis, and inference. The study included both generating hypotheses and testing these hypotheses in an iterative manner. The first wave of the study included interviews with 355 adults 65 years and older who have been identified for having high depression scores. Interviews were conducted using a fixed-choice response format. They followed up with in-depth semi-structured interviews of 102 individuals a year later. The sample of the second strand was selected purposefully to include the individuals with high- or low depression scores, along with high anxiety scores. Selection criteria were expanded in a next phase, to include a family history of depression, and also a demonstrated discordance between the severity of depression symptoms (scores on the Center for Epidemiology Studies Depression Scale or CES-D) and physicians’ opinion about the presence of depression, (i.e., when participants’ depression scores in the first strand were inconsistent with their physicians’ diagnoses).

Three types of data were collected in the second sample: (1) freelisting (asking each respondent to create a list of words that described a person who is depressed, as well as themselves when depressed); (2) semi-structured interviews including responses to vignettes about depression and open-ended questions; and (3) structured interview data from the first strand (a year ago), including demographic variables, scores of the CES-D and Beck Anxiety Inventory (BDI), and other scales.

Data analysis included a variety of iterative quantitative and qualitative strategies. Cultural consensus analysis was used for the analysis of freelist data, using the ANTHROPAC computer program. Cultural consensus analysis included the construction of lists of words/concepts that were also used by other members of the group with the same meaning. The frequency of mentioning each word and the average position of the word on the lists were calculated. The degree of similarity of each person to other members of the group was calculated by the data analysis program (ANTHROPAC) as an outcome of cultural consensus analysis. The obtained numerical indicators from these analyses were used in further statistical analysis.

Semi-structured interviews focused on answers to four broad questions about depression: “What is the cause of depression?” “What is it like to be depressed?” “What should you do for depression?” and “How would the depression turn out?” The responses were transcribed by an independent transcriber before being content-analyzed. First analysis included two team members who read each transcript and made an overall summary for that case, highlighting the compiled answers to each of the four questions mentioned above. Other group members assigned more refined codes to statements, leading to identification of (emergent) themes. These themes were then linked to the participants’ quantitative data/attributes.

In the last stage of these analyses, there was overlap between both types of data (themes and quantitative data) with closed-ended questions about loneliness (“How often did you feel lonely in the past week?”). Tests of association, group differences, and descriptive statistics were used to link the qualitative results to the variables assessed.
Data Collection in Mixed Methods
Psychological Studies

Mixed methods studies employ both qualitative and quantitative designs, data (or forms of data), and approaches to inquiry. Therefore, they have the flexibility of employing a wider variety of data collection strategies than do studies employing qualitative or quantitative studies alone. These strategies range from highly naturalistic and emergent/unstructured to highly molecular, structured, and closed-ended. The outcome is a combination of numerical, narrative, and graphic data. A part of the flexibility of mixed methods research is in being able to use multiple approaches within any data collection technique. For example, if data are collected via observations, highly structured protocols yielding numeric data may be used in combination with extensive field notes (or video/audio recordings) to capture both the occurrences and their context. As we discuss in the next section, the flexibility is also transferred to analysis, in which both statistical and content-analytic techniques may be employed to analyze the same data (qualitizing, quantitizing) or multiple sets of data collected via different strategies.

Similarly, interviews and questionnaires may be used to collect both types of data. Open-ended (qualitative) interviews or questionnaires provide qualitative data. Structured (closed-ended) interviews and questionnaires provide quantitative data. Often, in mixed methods studies, interviewers use a funnel-sequenced protocol starting from broad questions/probes and gradually narrow down to more focused specific (emerging or preplanned) questions (“What do you think about students’ substance abuse?”).

This is quite common when investigators collect data via focus groups by simultaneously combining interviews and observations while the group discussions are directed from general to specific. Mixed methods questionnaires, especially the Web/computerized ones, may follow the same pattern of general to specific, combined with open-ended items, thus providing multiple types of data.

Data Analysis in Mixed Methods Studies

As mentioned earlier, because multiple types of (narrative and numeric) data are collected in mixed methods studies, they must be analyzed accordingly. Numerical data are analyzed statistically in order to answer parts of the research question. Narrative data are analyzed via content analysis to provide answers to other aspects. These answers are then integrated toward a comprehensive answer to the research question(s) (Lieber & Weisner, 2010; Onwuegbuzie & Combs, 2010). Although the process is often presented in a linear manner, in practice integrative data analysis is often iterative. For example, inferences are made while each type of data is being analyzed (even while the data are being collected, in some cases), they are evaluated for credibility (i.e., Are there other plausible interpretations of the same result?), and are modified on the basis of these considerations.

Wesely’s (2010) study of language-learning motivation demonstrates how qualitative and quantitative data were collected and analyzed in order to answer emergent questions in a sequential mixed methods study. Wesely studied 131 graduates of five public elementary schools with language immersion programs. This sample consisted of all children/parents who responded to a questionnaire that was sent home (36% response rate). These children responded to a 40-item Likert-type questionnaire. Among those whose parents agreed to a follow-up interview, 33 were interviewed. The data collected in both strands were analyzed, using a six-stage process:

1. Preliminary data analysis during data collection: Taking informal notes during the 3 months of conducting 33 interviews, trying to detect patterns, and so on.
2. Initial qualitative analysis: Transcriptions and initial coding of the interviews, developing a list of inductive and deductive codes representing the themes.
3. Initial quantitative analysis: Statistical analyses, including reliability analysis, groups comparisons (t-test), and multiple regression/correlation, involving subscales of the questionnaire: “[using] qualitative methods to provide insight into the quantitative data, as I could directly
address the topics of the subscales, such as 'Attitudes toward learning the language,' with specific questions during the interviews. As a result, those motivational components emerged more readily from the qualitative data during coding" (p. 305).

4. Second iteration of qualitative theme analysis: After staying away from the qualitative data, reexamined the themes and codes in Stage 4, developing a refined and consistent set of concepts/themes from the interviews, and explored the connection of each with the previous literature.

5. Exploration of integrated findings: Examined and reinterpreted the obtained qualitative themes by comparing each with the results of statistical analysis; “also considered the literature in the fields of both learning motivation research and immersion education research. This led me to understand how to refine, combine, and organize the presentation of my findings. I made an effort not just to focus on areas where the two data sources were congruent; I also looked at areas where one data source revealed an important finding that was not there in the other source” (p. 308).

6. Planned mini case studies (not reported).

Statistical analysis revealed conflicting results, for example, in the relationship between attitudes and persistence in the program across schools. Expectation of less positive attitudes among the students who left the program was not confirmed, and the relationship varied across schools. Inconsistencies were partly explained with qualitative results, suggesting three different conceptualization of immersion programs among the students: “as a mysterious process, a grammar-focused experience, or (when successful) as a result of effort and work” (p. 308).

Qualitative results revealed many misunderstandings regarding the process of learning a language “often focusing more on the mysterious, automatic process that they associated with learning the language through their content courses or, alternatively, on the decontextualized grammar lessons that fit the more traditional definition of language class. These definitions of language learning might have confounded their responses to the ‘Attitudes toward Learning the Language’ subscales” (p. 308).

The Yoshikawa et al. (2008) description of data analysis possibilities in mixed methods research would further demonstrate the flexibility and potency of this type of iterative data analysis in mixed methods studies:

An important corollary to [the] distinction between qualitative and quantitative data and data analysis is that all four combinations of these two categorizations are possible. That is, qualitative data can be analyzed through either qualitative or quantitative data analysis techniques, as can quantitative data. Interview transcripts can be reliably coded for the frequency of mention of themes, the numbers of words or keywords, or the complexity of vocabulary and statistically analyzed. Ethnographic data from the world’s cultures have been coded for quantitative analysis . . . Conversely, individuals above or below a cut-off on a Likert scale or continuous dimension can be analyzed and characterized qualitatively, without further numeric representation. (p. 345)

Onwuegbuzie and Teddlie (2003) and Teddlie, Tashakkori, and Johnson (2008) have provided concise step-by-step strategies for data analysis in mixed methods. Some of these strategies are described below. Analyze qualitative and quantitative data independently, and integrate the findings at the end.

- Transform qualitative data to numeric, or quantitative data to qualitative profiles, graphic representations, themes, or groupings. Analyze each data accordingly (content analysis and/or statistical), link the findings, and make conclusions on the basis of both sets.
- Create groups of people/settings on the basis of qualitative data/observations, and then use statistical methods to compare the groups on quantitative data.
- Identify groups of attributes/themes through qualitative content analysis of narrative data, then use these identified attributes/obtained themes as variables in a subsequent statistical analysis. Make conclusions on the basis of both sets of findings.
- Form groups of people/settings on the basis of quantitative data, collect and analyze in-depth qualitative data to compare groups (e.g., to understand why they were different). Make conclusions on the basis of both sets of findings.
- Identify the components of the constructs under consideration among participants through factor analysis of quantitative data, then validate and confirm these obtained sub-constructs by collecting and analyzing detailed qualitative interviews/observations that are subjected to thematic analysis. The goal is to confirm, disconfirm, modify, or expand the meaning and nature of the “factors” obtained in the previous quantitative analysis.
- Use inherently mixed data analysis techniques (Teddlie & Tashakkori, 2009) that provide both qualitative and quantitative outcomes. For example, use social network analysis to obtain both graphic (qualitative) snapshots of communication networks and numeric indicators of various aspects of communication patterns.
Further developing these steps, Castro, Kellison, Boyd, and Kopak (2010) have suggested an integrative mixed methods (IMM) data analysis, using Atlas.ti, with six steps, from developing focus questions to creating story lines. Their proposed data analysis framework creates what they have called “thematic variables” from content analysis of narrative data, and links these variables to quantitative “measured variables” (p. 353). Examples are correlational analysis and exploratory factor analysis including both sets of variables.

Their last stage of analysis is “going full circle” by developing story lines that are anchored in narrative data and qualitative observations, which results in “recontextualizing” the results. This process involves linking the statistical results to “select indicated quotes to generate stories that ‘give voice’ to the very people who stated them” (p. 354). As an example, they compare and contrast narrative responses of five participants with highest life satisfaction scores to five with the lowest scores, in a “contrasting groups analysis” (p. 354, italic in the original). The final outcomes of this process are narrative profiles of these two groups, what they have called story lines: “Men who value and engage in family caretaking exhibit high levels of... positive machismo... in their male gender role identity, are giving and responsible, and they also experience high levels of life satisfaction” (p. 355, italics in the original).

Discussing the practical challenges in mixed methods research, Lieber and Weisner (2010) provide another conceptual framework for data analysis (and presentation). They provide examples of how fieldwork (field notes) can generate multiple forms of data, analyzed and integrated for comprehensive understanding of the phenomena/behaviors. Using the EthnoNotes software package as an example, they provide step-by-step illustrations for data analysis and integration. Figure 15.6 presents their conceptual model for mixed data analysis.

Making Inferences and Assessing Their Quality

Interpretation of the findings and making inferences are the most important aspects of any study. An interesting potential of mixed methods research is that mixing may occur at the meta-inference stage of the study, by trying to compare and contrast the findings of different strands. Torney-Purta’s (2009) suggestions for incorporating the cultural context in interpretation of quantitative results might be an example of such integration (see earlier).

Traditionally, most discussions of quality in psychological research have focused on internal validity and variations of it (statistical conclusion validity) and
generalizability (external validity, construct validity). Qualitative researchers have questioned the adequacy of these conceptualizations of quality, and have proposed alternatives to them, focusing largely on concepts of trustworthiness and credibility.

Mixed methods researchers must take these diverse approaches to quality into consideration. Both the inferences that are made in the qualitative and quantitative strands and the integrated meta-inferences must be systematically examined for quality (O’Cathain, 2010). Evaluating the possibility of alternative plausible interpretations of the findings of each strand of a research/evaluation project is necessary in all research, but this is especially the case with mixed methods.

Although at the first glance, having to use two (or multiple) sets of standards of quality might seem impossible, closer attention would reveal that this is not the case for various reasons. Among them is the fact that inferences from each set of results (i.e., from qualitative or quantitative analyses) may be evaluated separately. A second is that there are major similarities between the alternative sets of audits/standards, at least in the underlying concepts and/or assumptions (Camic, Rhodes, & Yardley, 2003). The first (separate evaluation of quality for different sets of results) might be problematic for highly complex designs in which the two approaches to design, data collection, data analysis, and interpretation might iteratively influence each other on an ongoing basis (e.g., results of one phase of the study might influence the data analysis outcomes of the other). Regardless, familiarity with the two (or multiple) approaches to quality might potentially lead to creating bridges between these audits/standards, or might lead to comparing/contrasting them in a dialectic manner (Greene & Hall, 2010).

We (Tashakkori & Teddlie, 2003, 2008) have proposed examining two broad domains of inference for the evaluation of quality. One is the degree to which the researcher has planned and effectively implemented all the necessary procedures for finding answers to his/her research questions (design quality). The other is the degree of adequacy/plausibility of the conclusions following each of the sets of findings (interpretive rigor). We have kept the issues of transferability/utility separate but highly dependent on inference quality.

An important component of interpretive rigor—perhaps the most important for mixed methods researchers—is the evaluation of the adequacy of integration (integrative efficacy): the degree to which a mixed methods researcher has effectively integrated the findings and conclusions of various strands of the study, to provide comprehensive answers to the research questions. Effective integration requires the mixed methods researcher to connect the inferences that are gleaned from each strand of the study, evaluate their similarities and inconsistencies, and provide credible explanations for these variations and similarities. In that process, the inferences of one strand may be used to elaborate, modify, understand, confirm, or provide the limits of the inferences gleaned from the other strand.

Despite the fact that consistency between the inferences derived from qualitative and quantitative strands is often considered desirable, multiple plausible interpretations/inferences, or complementary findings, occur frequently in mixed research and are considered to be valuable outcomes. Plausible inconsistency or difference (i.e., those that are not deemed to be due to problems in design, data, or analysis) might indicate the presence of equally credible explanations for the phenomenon or relationship, or multiple (e.g., complementary) components/aspects of the phenomenon or construct. These findings can also have great heuristic value for future research.

A final note on the issue of quality is related to outcomes of research. Elsewhere (Tashakkori & Teddlie, 2010a), we have pointed to the necessity of assessing the utilization quality of mixed methods findings and recommendation. This necessity may be demonstrated with two questions:

1. To what extent has using mixed methods provided more credible explanations and understandings (what O’Cathain, Murphy, & Nicholl, 2007 have called “mixed methods yield”)?
2. How useful are the explanations and recommendations for policy and practice planning (i.e., utilization quality)?

The first is closely related to integrative efficacy. If the findings of various strands of a mixed methods study are not well integrated, utilizing mixed methods does not yield more credible inferences than disparate approaches/methods.
Perhaps one of the reasons for the relative popularity of mixed methods in recent years has been rooted in the assumption of increased credibility. Evidence of this assumption might be found in agency requests for proposals (RFPs), increase in the number of grants utilizing mixed methods (Dahlberg, Wittink, & Gallo, 2010; Plano Clark, 2007; Song et al., 2010), and the number of papers accepted for publication in journals. Song et al. provide detailed suggestions for preparing such grant proposals as well as manuscripts for publication.

The second question (impact on policy and practice) depends heavily on the scope and generalizability of the results and creates a bridge between inference quality and inference transferability. This issue is probably more controversial than the first one, in the context of the evidence-based practice movement (and randomized control trials) of recent years (Song et al., 2010). Our answer to this question assumes that a desirable (mixed methods) research finding (“research that matters,” Torney-Purta, 2009, p. 857) is the one that has good potential to impact policy, planning, and intervention. We have called this the utilization quality of inferences, which links the quality (trustworthiness, credibility, validity) and transferability of integrated inferences. Within this framework, assessing the quality of recommendations becomes a crucial quality audit in mixed methods psychological research.

A common assumption in employing mixed/multimethods approaches to research/evaluation is that, because of their potential for broader understanding of social issues, they provide more robust opportunities for devising policies and practices to implement positive change (Greene, 2011). It is essential to assess the degree to which a mixed research/evaluation project has the potential to offer credible policy and practice recommendations, and even the degree to which these outputs impact decisions. This value-added policy and practice quality has been assumed almost as a truism in many large-scale projects, and also in calls for proposals, but remains to be systematically demonstrated in any given research/evaluation project (Tashakkori & Teddlie, 2010a).

**EPILOGUE: ONE WORLD, MANY LENSES**

Issues discussed in this last section are particularly important for the more traditional social/behavioral sciences (i.e., those with the longest histories and the most well-defined methodologies). Psychology certainly belongs in this category since many of our concepts regarding how to conduct, analyze, and interpret findings originated from renowned scholars in the field. While qualitative and mixed methods are more acceptable now, there is still widespread agreement with the assertion of Cook and Campbell (1979) that:

> We assume that readers believe that causal inference is important and that experimentation is one of the most useful, if not the most useful, way of gaining knowledge about cause. (p. 91, italics in original)

It can be safely concluded that many academic departments in psychology focus on quantitative methods and many psychology journals still favor quantitative-only studies.

Much debate has occurred throughout the social/behavioral sciences over the past two decades about the feasibility of mixed methods research. Among the perceived obstacles, scholars have pointed to the impossibility of taking multiple perspectives, the difficulty of acquiring both sets of methodological skills, the disparity of language, and the inconsistency of quality standards. Most of these identified obstacles have been successfully dealt with in numerous projects, in different ways. We offered examples of such projects above. Many more may be found in the literature.

Perhaps the most unresolved and contentious current controversy revolves around the issue of multiplicity of paradigms/worldviews. A full discussion of these controversies is beyond this chapter, and may be found in other sources (e.g., Denzin & Lincoln, 2011; Johnson & Gray, 2010). Our reaction to these criticisms is rooted in our own training as psychologists and in our experiences in conducting research in complex human settings.

We consider mixed methods a natural extension of the “naïve researchers” process of answering questions, more so than either of the qualitative or quantitative approaches alone (Tashakkori & Teddlie, 2010b). This humanistic image of the researcher makes room for simultaneous and/or sequential utilization of “qualitative” (emic-level, subjective, emergent, arts-based, etc.) and “quantitative” (etnic-level, objective, structured, etc.) data and operations. Within such a framework, the multiplicity of worldviews does not pose a problem, since they are employed side by side in a dialectic manner (Greene, 2007; Greene & Hall, 2010), weighed against each other, or are used at different times. The human researcher does not lose this capability when taking the role of formal researcher, since mixed methodologists are free to utilize the whole spectrum of tools (from qualitative and quantitative communities/approaches) for collecting and analyzing data to answer research questions. In Gorard’s (2010) words:

> No one, on buying a house, refuses to discuss or even know the price, the mortgage repayments, the room measurements
or the number of bathrooms. No one, on buying a house, refuses to visit the house, look at pictures of it, walk or drive around the neighborhood or talk to people about it. All rational actors putting a substantial personal investment in their own house would naturally and without any consideration of paradigms, epistemology, identity or mixed methods, use all and any convenient data to help make up their mind.…. We would not refuse to visit the house, or talk to the neighbors about it, because we were “quantitative” researchers and did not believe that observation or narratives were valid or reliable enough for our purposes. We would not refuse to consider the interest rate for the loan, or the size of the monthly repayments, because we were “qualitative” researchers and did not believe that numbers could do justice to the social world.…. We collect all and any evidence available to us as time and resources allow, and then synthesize it naturally, without consideration of mixing methods as such. (p. 246)

Although, in principle, mixed methods approach has strong roots in behavioral methodology, it is also a new way of doing research, within the context of advances in qualitative and quantitative methods/approaches. It provides the researcher with the insight and the capability to utilize all necessary methods of inquiry for answering his or her research questions that need such multiplicity. Questions that simultaneously ask what, why, and how are prime examples of this need (Plano Clark & Badiee, 2010). Given the prevalence of these three questions in psychological inquiry, and their strong interdependence in many psychological research projects, mixed methods might provide more effective investigative tools than the isolated qualitative or quantitative approaches.

REFERENCES


Queries in Chapter 15

Q1. Please provide the reference for “Davidson & Stayner, 1997”
Q2. Please provide the reference for “Miller, 1994”
Q3. AU: Please confirm Cialdini change from 1978 to 1976 per Refs
Q4. Please provide the reference for “Calves, Cornwell, and Envegue, 1996”
Q5. AU: Please confirm years of McGuire entries per Refs
Q6. Please provide the reference for “Newman & Benz, 1999”
Q7. Please provide the reference for “Barg et al. (2008)”
Q8. Please provide the reference for “Tashakkori & Teddlie, 2003”
Q9. Please provide the reference for “Plano Clark, 2007”