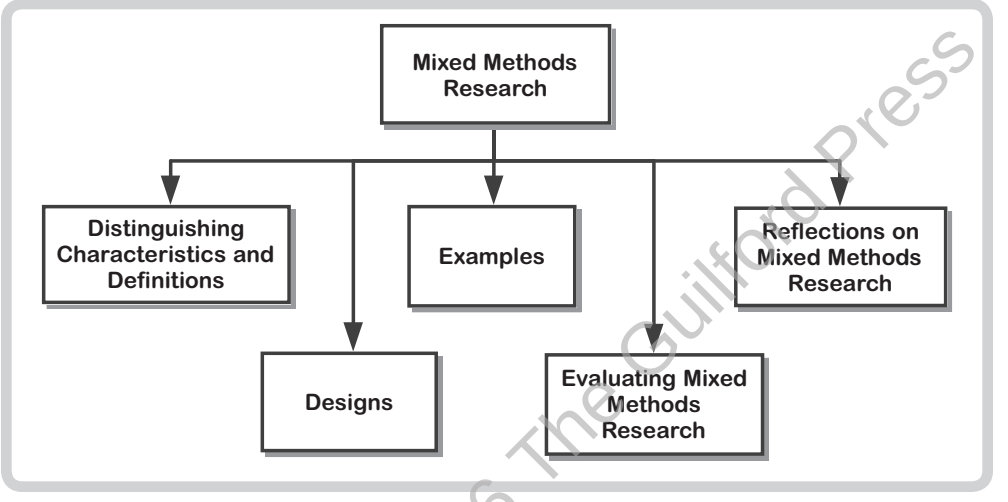
CHAPTER THIRTEEN

Mixed Methods Research



Mixed methods research combines traditional qualitative and quantitative methods to answer questions that neither method can answer on its own.

CHAPTER OBJECTIVES

- ✓ Understand the purpose and premise underlying mixed methods research.
- ✓ Understand how mixed methods research combines quantitative and qualitative methods.
- ✓ Understand the difference between concurrent and sequential mixed methods designs.
- ✓ Understand triangulation and embedded procedures for concurrent mixed methods research.
- ✓ Understand explanatory and exploratory procedures for sequential mixed methods research.

DISTINGUISHING CHARACTERISTICS AND DEFINITIONS

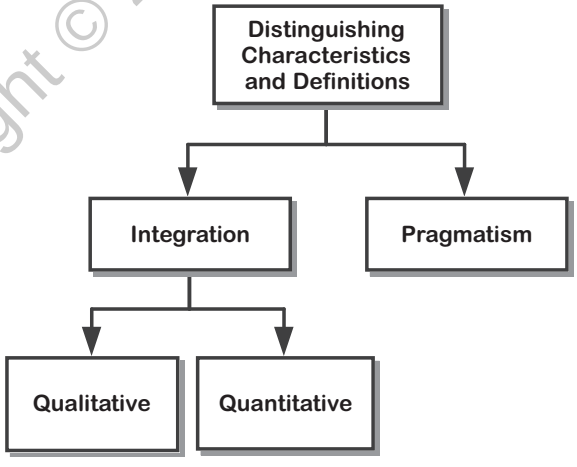
MAP 13.1

Mixed methods research combines and integrates quantitative and qualitative methods in a single study.

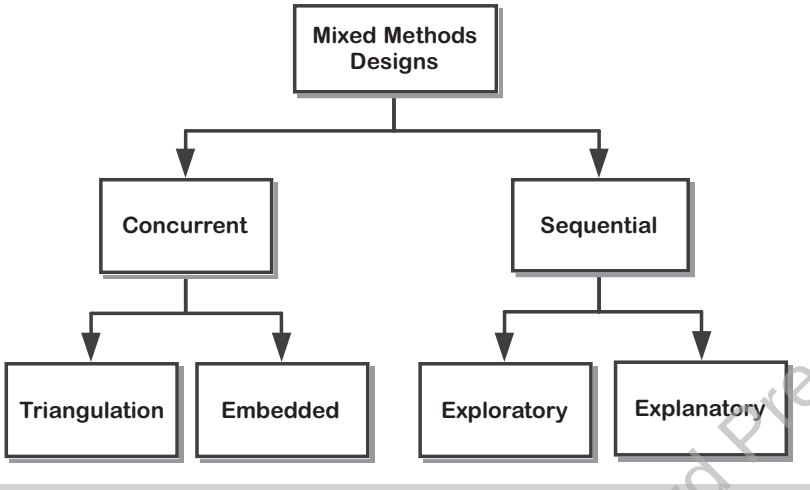
The purpose of mixed methods research is to investigate a problem fully by drawing on quantitative measures to determine frequencies and relationship of variables, as well as on qualitative tools to provide insight into meaning and understanding. It combines qualitative and quantitative methods in a way that emphasizes the strengths of each method and avoids overlapping weaknesses.

> It involves the recognition that all methods have their limitations as well as their strengths. The fundamental principle is followed for at least three reasons: (a) to obtain convergence or corroboration of findings, (b) to eliminate or minimize key plausible alternative explanations for conclusions drawn from the research data, and (c) to elucidate the divergent aspects of a phenomenon. (Johnson & Turner, 2003, p. 299)

The philosophy that undergirds mixed methods research is *pragmatism*, which is a quintessentially American philosophy advocated by William James, Charles S. Peirce, and John Dewey. Pragmatists propose that the value of an inquiry can best be judged by its practical consequences. Mixed methods research fits the pragmatic idea because it makes practical use of both induction and deduction to achieve understanding and explanation (Johnson & Onwuegbuzie, 2004, p. 14).



MAP 13.1



MAP 13.2

MIXED METHODS DESIGNS

MAP 13.2

Some mixed methods studies weight the qualitative and quantitative strands equally, while others weight one strand more heavily than another. It is important to distinguish mixed methods designs from multiple-methods research that uses more than one method of data collection and analysis within the same research tradition. For example, ethnographies and case studies use interviews, observations, and documents/discourses. Similarly, quantitative studies may depend on both surveys and measures of academic achievement. A unique feature of mixed methods research is that qualitative and quantitative data are separately collected and analyzed, and are then brought together in a final interpretation, in what are known as *metainferences* or *integrated mixed inferences* (Tashakkori & Teddlie, 2003a). There are two basic types of mixed methods designs: *concurrent* and *sequential*, each with specific procedures for data collection and analysis.

Concurrent Designs

Concurrent mixed methods designs "are those in which the researcher converges or merges quantitative and qualitative data in order to provide a comprehensive analysis of the research problem" (Creswell, 2009, p. 228). There are two procedures for data collection and analysis in concurrent designs: *triangulation* and *embedded*.

1. In the *concurrent triangulation design*, there is one data collection phase in which the qualitative and quantitative data are collected simultaneously, and there are two separate analyses of the qualitative and quantitative data. In the interpretation phase, either the findings are merged, or they are compared in a discussion section. In this procedure, the qualitative and quantitative strands are usually weighted equally. Figure 13.1 illustrates this design.

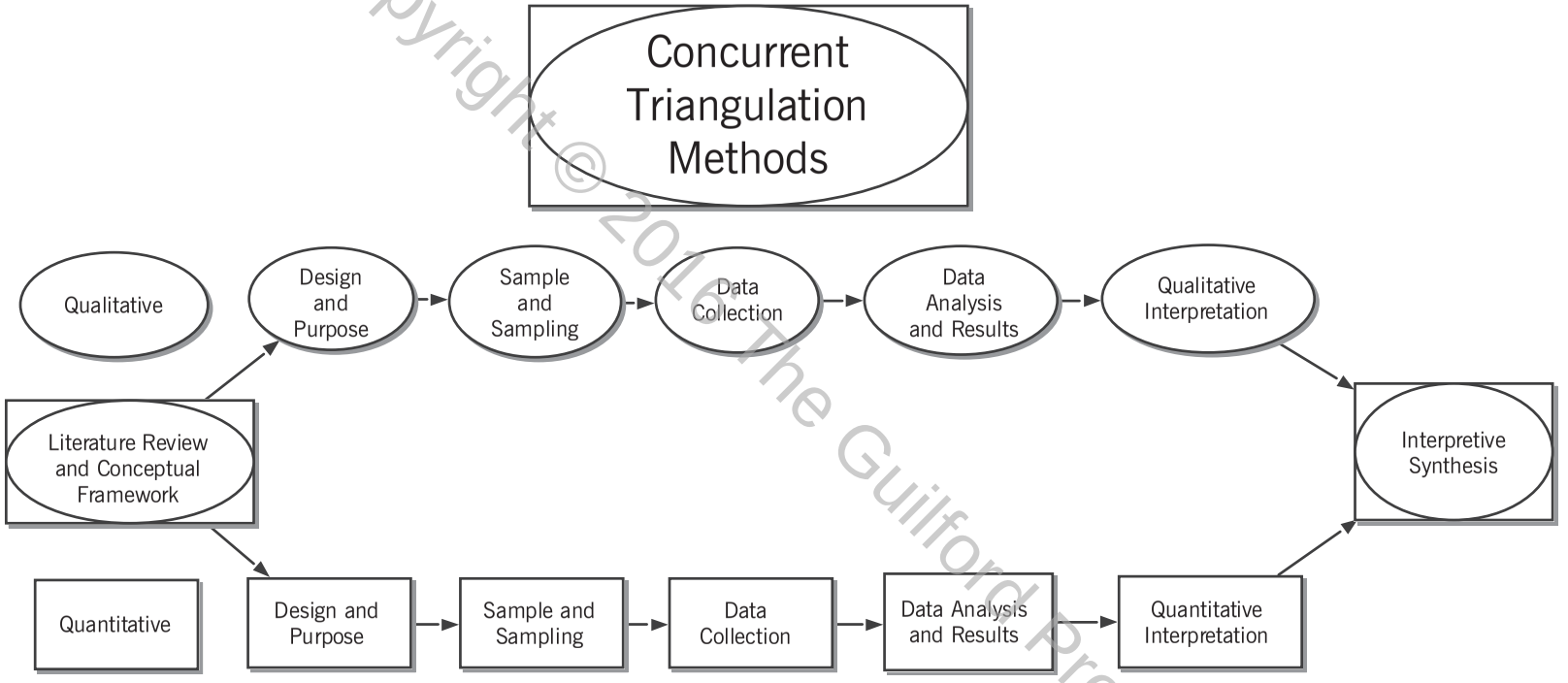


FIGURE 13.1. Concurrent triangulation procedures.

2. In the *concurrent embedded design*, the procedure for data collection is similar to that in the triangulation design: There is one data collection phase in which qualitative and quantitative data are collected simultaneously. However, during the data analysis phase one strand is nested within another stage, which is more heavily weighted. Figure 13.2 illustrates this design.

Sequential Designs

Sequential mixed methods designs "are those in which the researcher seeks to elaborate on or expand on the findings of one method with another method" (Creswell, 2009, p. 234). Within sequential designs, there are also two procedures for data collection and analysis: *explanatory* and *exploratory*.

1. In the *sequential exploratory design*, the qualitative strand is weighted more heavily. The quantitative strand is used to assist in interpreting the qualitative findings. Figure 13.3 illustrates this design.
2. In the *sequential explanatory design*, the quantitative strand is weighted more heavily and informs procedures in the qualitative strand. The qualitative analysis is used to examine or clarify quantitative findings. Figure 13.4 illustrates this design.

EXAMPLES OF MIXED METHODS DESIGNS

1. Concurrent Embedded Design

Feldon, D., & Kefai, Y. (2008). Mixed methods for mixed reality: Understanding users' avatar activities in Virtual Worlds. *Educational Technology Research and Development, 56*(5–6), 575–593.

This study investigated the use of avatars in a game called Virtual Worlds, and it used several methods of data collection and analysis. There were 595 participants in the game, who created avatars that had to survive and progress in their environment, respond to other avatars, and deal with disease and health risks. Over 33% of participants in the game engaged in avatar activities through computer clicks, hits, and navigations. An avatar-related activity might be a change in physical appearance or the exchange of facial features as a trade or a symbol of friendship. The participants visited over 6.93 million screen locations over the 6-month duration of the study.

The authors used log data, online and offline observations, interviews, and surveys as data sources. The logs represented the raw data of participants' actions—the total count of clicks and hits. A 30-item online survey administered after the game included an outbreak of a smallpox virus, which produced spots on the face of each avatar. All but one of the survey questions were closed-response questions that focused on general use and user preferences. Interviews of 35 participants were conducted at the end of the study and consisted of questions such as "How is your avatar like you and/or not like you?" and "How often do you change your avatar?"

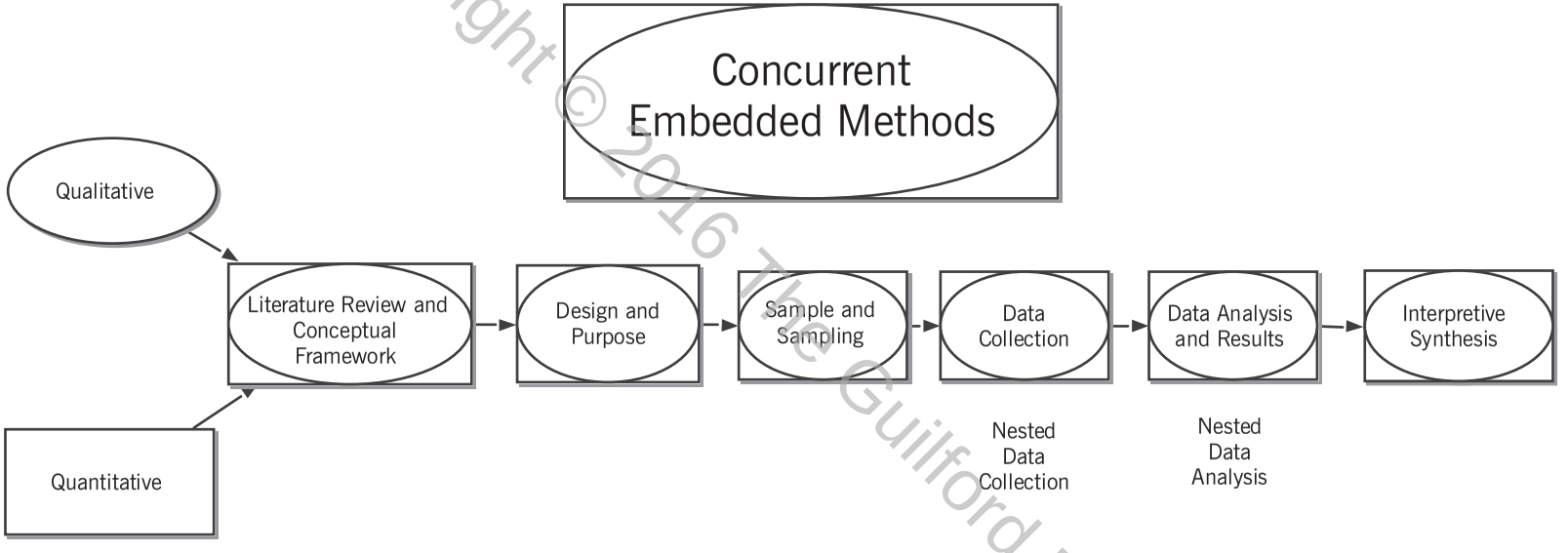


FIGURE 13.2. Concurrent embedded procedures.

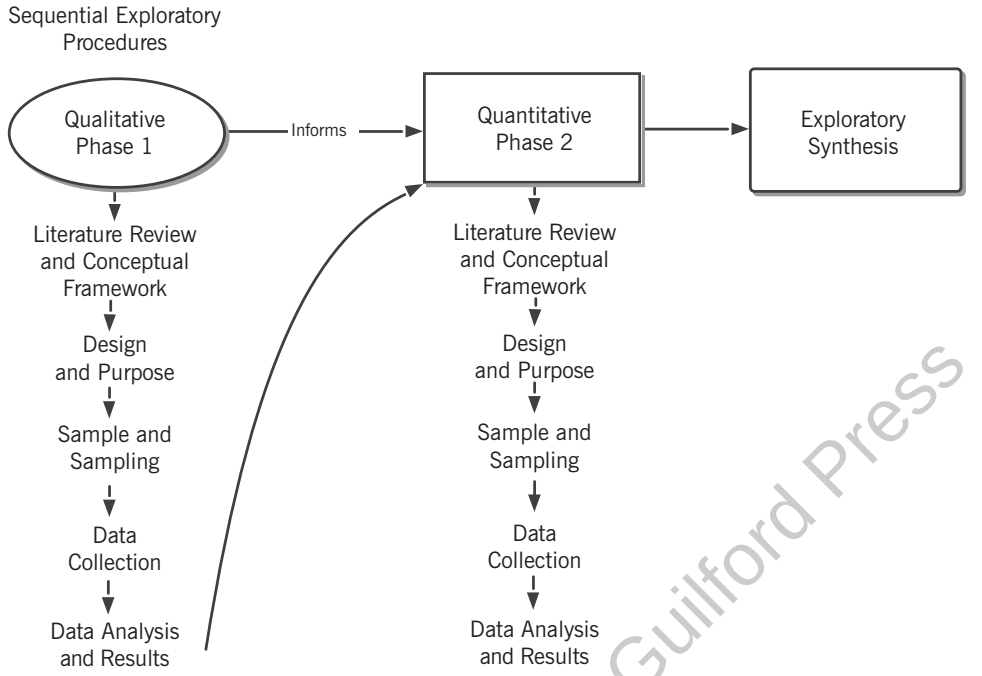


FIGURE 13.3. Sequential exploratory procedures.

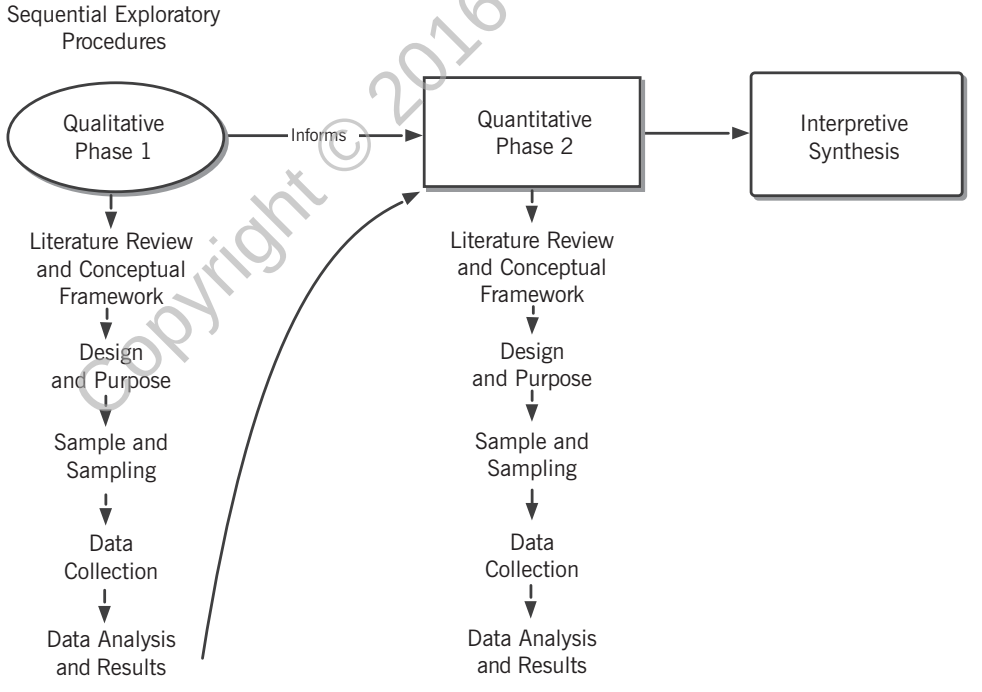


FIGURE 13.4. Sequential explanatory procedures.

(p. 583). A researcher, who was embedded in the game as a reporter, collected the observations in the online environment. The researcher visited virtual locations to observe the movements and interactions of avatars.

Data from the server logs and the surveys were analyzed, and three types of participants/users were identified: casual users, social users, and heavy users. This categorization formed the basis for analyzing differences across the total sample, and it provided a key quantitative comparison for the overall time and effort spent on avatar-related activities. The comparison of the three groups used an ANOVA, which showed that participants from the three groups were very similar in their avatar-related activities.

The combination of methods in this study provided an understanding of the incidence of participants' activities, the details of their motivation and questions, and their concerns about their virtual experiences. The researchers attempted to balance quantitative and qualitative methods and designs. There was a progression of methods, beginning with server logs and quantitative method, and ending with interviews and observations. This study is thus a good example of a concurrent embedded design.

2. Sequential Explanatory Design

Gasiewski, J., Eagan, M., Garcia, G., Hurtado, S., & Chang, M. (2012). From gatekeeping to engagement: A multi-contextual, mixed method study of student academic engagement in introductory STEM courses. *Research in Higher Education, 53*(2), 229–261.

This study "employed a sequential, explanatory mixed method approach to provide a richer understanding of the relationship between student engagement and introductory science instruction" (p. 229). The authors explained, "With this research design, we sought not only to examine the predictive power of specific learning strategies and classroom contexts that relate to STEM students' engagement in introductory courses but also to further support and enrich these findings through students' narrative experiences of being enrolled in these courses" (p. 230). The researchers reviewed research on academic engagement, active learning pedagogies, motivation, and faculty behavior to establish a theoretical framework for the study.

The more heavily quantitative strand was conducted first. The sample was drawn from 73 introductory science, technology, engineering, and mathematics (STEM) courses from 15 colleges and universities. The researchers administered surveys at the beginning and end of the courses; 2,873 students completed both surveys. The DV was academic engagement, which included these eight factors: "frequency with which students asked questions in class, discussed course grades or assignments with the instructor, attended professor's office hours, participated in class discussions, tutored other students in their introductory STEM course, reviewed class material before it was covered, attended review or help sessions to enhance understanding of course content, and studied with students from their introductory STEM course" (p. 237). An extensive multivariate analysis "suggested that 3.1 and 4.1% of the

variance in academic engagement was attributable to differences across classrooms and institutions, respectively. In other words, classrooms and institutions appear to have a marginal effect on students' academic engagement, and the vast majority of variance we see in academic engagement can be attributed to differences between students" (p. 239).

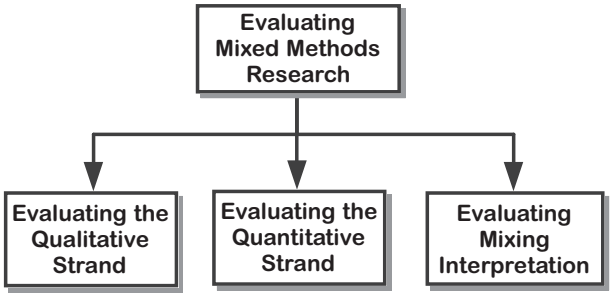
The qualitative strand used a purposeful, criterion sample of 8 colleges and universities selected from the original 15, based on survey responses and evidence of innovation in teaching practices. The researchers conducted 41 focus groups with students who had completed the quantitative surveys or who were currently enrolled in an introductory STEM course. A constant comparative strategy was used to code and analyze data. In the final step of analysis and interpretation, the researchers combined findings from the quantitative and qualitative strands. Though the statistical analysis yielded no significant connection between student engagement and teaching practice, it did provide evidence about the relationship between student attributes such as excitement about learning, competitiveness, and career orientation on the one hand, and engagement and success in the courses on the other. Interviews with students supported this connection and provided insights that furthered understanding.

The researchers integrated the quantitative and qualitative findings to develop two composite types of STEM instructors: *gatekeeper* and *engaging* professors. Gatekeeper professors "disregard individual learning styles because they are so focused on conveying the abundance of information that must be passed on to students who are worthy of passing through the gates. Their expectation is that students can and should understand the content at a sophisticated level" (p. 252). By contrast, an engaging professor "uses strategies that encourage active learning, cooperation among students, and student–faculty contact. . . . facilitates student excitement in the classroom through humor, enthusiasm, and practical application . . . is highly accessible to students and encourages them to participate in additional learning opportunities offered by the university" (p. 253).

The researchers concluded, "If educators are the key change agents in this dynamic, the findings suggest that introductory STEM course instructors must think just as carefully and thoroughly about how they interact with and come across to students as they do about the course content and how to assess its mastery, especially when it comes to scaling up STEM achievement and increasing student persistence" (p. 256).

EVALUATING MIXED METHODS RESEARCH MAP 13.3

The evaluation of mixed methods studies builds on the categories and criteria for qualitative and quantitative research, and adds a third consideration: the mixing or interpretation of the methods. Accordingly, there are three steps in evaluating mixed methods studies:



MAP 13.3

1. Evaluating the qualitative strand. The criteria for this strand are trustworthiness and transferability.
2. Evaluating the quantitative strand. The criteria for this strand are theory or framing construct, sample and sampling, and data collection and analysis.
3. Evaluating mixing/interpretation: Is there a clear description of how the data were mixed in the study? Were they mixed at the sampling, data collection, and/or data analysis (interpretation) stage? This step involves an evaluation of two elements:
   - Timing: What was the timeline of the study in terms of the sampling, data collection, and data analysis? Does the article include a timeline that visually depicts the timing of each step?
   - Weight: What was the emphasis on each strand? Did the study emphasize qualitative and quantitative methods equally? Or was there a clearly communicated emphasis on one strand or the other?

The rubric in Figure 13.5 provides the criteria for evaluating mixed methods studies.

Example: Evaluation of the Virtual Worlds Study

Qualitative Strand

- **Trustworthiness:** The Virtual Worlds study (Feldon & Kefai, 2008) merits a moderate to strong rating for trustworthiness. There is a thorough literature review, together with lengthy descriptions of the researchers' involvement in the study and of the online and offline environments. The study included triangulation of sources of data as well as a team of researchers to check for accuracy. A variety of qualitative data were collected and analyzed, both online and offline, so that users' perspectives were represented well. The article includes little or no discussion of coding and development of themes.

Qualitative strand	**Strong**	**Moderate**	**Weak**
Trustworthiness	Thick description and quotes provided; purposive or theoretical sampling described; explicit triangulation; checks for accuracy; clear procedures for coding and development of themes; sufficient engagement to earn trust.	Some thick description and quotes provided; sampling discussed; triangulation implied; limited checks for accuracy; partial description of procedures for coding and development of themes; sufficient engagement to earn trust.	Limited or no thick description and quotes provided; scant mention of sampling; triangulation not in evidence; no reported checks for accuracy; limited (if any) description of procedures for coding and development of themes; insufficient engagement.
Transferability	Detailed description of context and actors' actions, thoughts, verbatim language from interviews and documents.	Partial description of context and actors' actions, thoughts, verbatim language from interviews and documents.	Limited description of context and actors' actions, thoughts, verbatim language from interviews and documents.
Quantitative strand	**Strong**	**Moderate**	**Weak**
Theory or framing construct	Well-developed literature review that establishes either a theory or a framing construct and includes 10 or more references; strong support for the IV and DV; clear identification of hypothesis or research questions and the DV.	Literature review may not adequately develop a theory or framing construct, or includes fewer than 10 references; research questions implied; adequate support for the IV and DV.	Unclear literature review that does not develop a theory or framing construct; fewer than five references; no research questions evident; inadequate support for the IV and DV.
Sample and sampling	Clear description of the sample characteristics and the population; clear description of sample; evidence of sufficient size (60 or more); random selection of sample or sample matching.	Description of sample is present, but lacking details about characteristics and the population; scant detail about variability of sample; insufficient sample size.	No description of the sample; total lack of details about the population; sample neither randomly selected nor matched.
Data collection and analysis	Clear description of measures selected to collect data on DV and evidence of validity and reliability (*r*); correct use of tests for DV clear; description of statistical significance (*p*) and alpha level, as well as of the inferential tests used; researcher stays close to data and does not overconclude; article includes explanatory charts and narratives.	Mention of measurement, but lacking in details about validity and reliability (*r*); mention of statistical significance (*p*) without including inferential tests used; researcher avoids overconcluding; article includes explanatory charts and/or narratives.	Minimal or no mention of measurement; no mention of validity or reliability; no mention of significance; researcher may overconclude; article does not include adequate visual or narrative explanations.

(continued)

FIGURE 13.5. Rubric for evaluating mixed methods studies.

Mixing/ Interpretation	Strong	Moderate	Weak
Timing	Clear and well-developed rationale for mixing data; article identifies organization of the design as concurrent or sequential; clear visual representation of the timeline–methods mix (timeline could stand on its own).	Mention of strategies for mixing data, but incomplete description; article may be unclear whether design is concurrent or sequential; unclear or missing visual representation of timeline–methods mix.	No mention of strategies for mixing data; no mention of concurrent or sequential design; missing visual representation of timeline–methods mix.
Weight	Clear and well-organized discussion about the purpose of the study and how it affects the weighting of the data.	Mention of the purpose of the study, but insufficient explanation of how it affects the weighting of the data.	No mention of the purpose or how the purpose affects the weighting of the data.
Overall rating	**Strong**	**Moderate**	**Weak**

FIGURE 13.5. *(continued)*

• **Transferability:** The study's transferability is strong. There are detailed descriptions of the context, both the online and offline environments. Specific quotes from field notes, videotaped interviews, and observations depict the actors' action and thoughts. In addition, the article includes photographs of avatars and the gaming "dashboard" to engage the readers.

Quantitative Strand

• **Theory or Framing Construct:** The Virtual Worlds study merits a moderate rating for theory or framing construct. The study examined online avatar-related activities, but the article provides little discussion about the value of online environments for learning or leisure. Rather, the study appeared to focus more on the use of mixed methods and an effort to address the shortcomings of a quantitative study based on the use of surveys and server data.

• **Sample and Sampling:** The study's sample and sampling are strong. The sample was large enough for a nonexperimental design based on correlations, with 595 children responding to the survey, and a subset of 88 students responding to additional survey questions. Approximately 70 million lines of server log data were analyzed as well.

• **Data Collection and Analysis:** The study's data collection and analysis also merit a strong rating. As mentioned earlier, there was a nonexperimental group comparison analysis of casual, social, and heavy users with server log data, including time spent on activities in a variety of online gaming locations. The 30-item online surveys used for pre- and postactivity analysis are described in detail, and the overall Cronbach's alpha reliability coefficient was .72.

Mixing/Interpretation

- **Timing:** The study earns a moderate rating for timing. The researchers used a framework for the mixing and analysis of data: expansion, triangulation, complementarity, initiation, and development. However, the article includes no explicit timeline or visual representation of the mixed methods.

- **Weight:** The study also earns a moderate rating for weight. From the authors' description of the study, it seems that their intent was to have an equal weighting of qualitative and quantitative data. The purpose of collecting the qualitative data was to provide the researchers with understanding of the gaming culture and the motives of users and their avatars, but it is not made clear how this was enacted.

Example: Evaluation of the STEM Course Study

Qualitative Strand

- **Trustworthiness:** The Gasiewski et al. (2012) study merits a strong rating for trustworthiness. The literature review of approximately 100 references establishes a very good rationale for the study of STEM courses in higher education. The study used a sequential explanatory design with a purposeful sampling strategy. The large volume of focus groups, with 2–10 students per group, provided triangulated data and increased the likelihood of data saturation. The researchers' categories are illustrated by numerous quotes and descriptions; intercoder reliability was established through multiple steps to strengthen the integrity of data.

- **Transferability:** The study's transferability is also strong. Based on the interviews, the authors created composite representations of the professors and students that were not identifiable. The representations were developed with numerous quotes and observations and in-depth descriptions of college classrooms and campuses.

Quantitative Strand

- **Theory or Framing Construct:** The study merits a strong rating for theory or framing construct. The literature review is current and includes approximately 100 references, discussing both the questions about engagement in STEM learning and the use of mixed methods research designs for investigations. The DV was student engagement (including persistence and academic performance in STEM courses); the IVs of interest were quality of teaching and the learning environment (in particular, the items on the behavioral academic inventory). The research questions are clearly stated and match the study's purposes.

- **Sample and Sampling:** The study's sample and sampling are also strong. The sample was large enough for a nonexperimental correlations design, with 15 higher education institutions and 73 classrooms. Although the sample was nonrandom, a total of 3,205 students filled out the surveys; of those, 2,873 students were included in the data analysis.

• **Data Collection and Analysis:** The study likewise earns a strong rating for data collection and analysis. Survey data were collected from both students and college professors. The behavioral academic inventory was developed from an earlier survey on the development of scientific dispositions and collegiate habits of mind. Factor analysis of this survey helped to ensure construct validity, with the identification of multiple factors used as both IVs and DVs. Reliability for survey was established by using Cronbach's alpha; for example, the factor for academic engagement had a Cronbach's alpha of .75. Extensive techniques were used to weight the data and to account for missing data in the hierarchical linear model statistical analysis.

Mixing/Interpretation

• **Timing:** The study earns a strong rating for timing. The design of the study was identified as a sequential explanatory design. The rationale for timing and mixing qualitative and quantitative methods is made very clear in the article and is presented in a visual model (a flow chart of the mixed methods design procedures).

• **Weight:** The study merits a moderate to strong rating for weight. The purpose of the study was to use the qualitative data to explain the quantitative findings, thus separating the study into two major phases. The article makes no mention of the precise weighting of the data, however.

REFLECTIONS ON MIXED METHODS RESEARCH

Mixed methods research is not without its critics. Methodological purists hold that research is either qualitative or quantitative, and that nothing can exist in between. Despite these reservations, most researchers acknowledge mixed methods research as a promising development. The publication of the *Handbook of Mixed Methods in Social and Behavioral Research* (Tashakkori & Teddlie, 2003b) added to its legitimacy, as did the founding of the *Journal of Mixed Methods Research*. Further evidence of its entry into the mainstream is establishment of a special interest group within the American Education Research Association that has the following as its goal: "To support, encourage, and increase dialogue and idea exchange among educational researchers utilizing mixed methods and those interested in integrating qualitative and quantitative research approaches" (see *www.aera.net/SIG158/MixedMethodsResearchSIG158/tabid/12201/Default.aspx*).

However, there are some lingering concerns about mixed methods approaches. They take longer to implement, entail more resources, require separate inductive and deductive analyses, and add a step to articulating qualitative and quantitative results. After interviewing 20 social scientists who had conducted mixed methods studies, Bryman (2007a) concluded that "insufficient attention has been paid to the writing up of mixed methods findings, and in particular to the ways such findings can be integrated. Indeed, it could be argued that there is still considerable uncertainty concerning what it means to integrate findings in mixed methods research" (p. 22).

CHAPTER SUMMARY

MAP 13.4

- ✓ Mixed methods research combines qualitative and quantitative methods of data collection and analysis.
- ✓ The two mixed methods designs are concurrent and sequential.
- ✓ Concurrent mixed methods designs combine or merge quantitative and qualitative data.
- ✓ Sequential designs expand on the findings of one method with another method.
- ✓ Mixed methods researchers combine findings in a final interpretation, called *metainferences/integrated mixed inferences*.
- ✓ Mixed methods research is evaluated according to the criteria for qualitative and quantitative studies and also according to two criteria for mixing/interpretation (i.e., timing and weight).

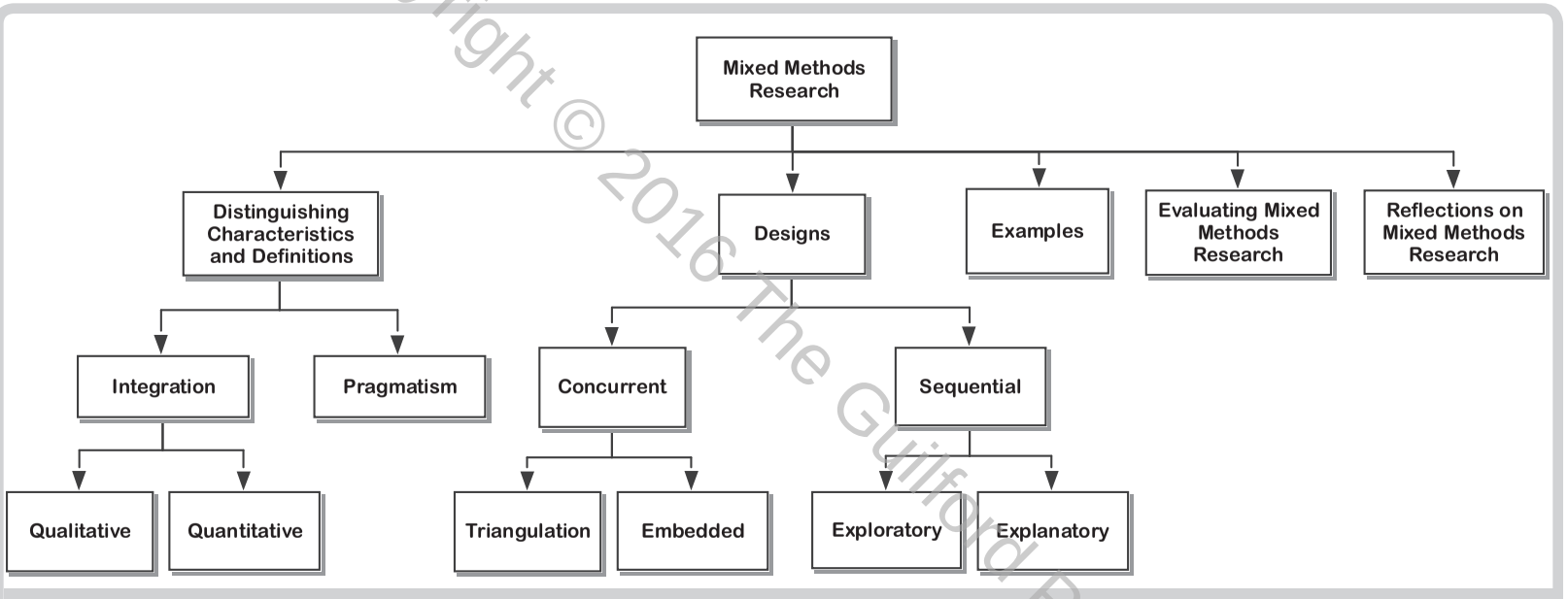
KEY TERMS AND CONCEPTS

concurrent embedded methods
concurrent designs
concurrent triangulation methods
metainferences/integrated mixed inferences
mixed methods research
multiple-methods research
pragmatism
sequential designs
sequential explanatory methods
sequential exploratory methods

REVIEW, CONSOLIDATION, AND EXTENSION OF KNOWLEDGE

1. Using an electronic database or a search engine, locate a mixed methods study on a topic of interest. Read the article, and then answer the questions below.
   a. What research is reviewed? Does it provide a rationale for the study? Does it provide a rationale for using mixed methods?
   b. What was the purpose of the study?
   c. What design was employed? Concurrent (triangulation or embedded) or sequential (explanatory or exploratory)?
   d. How were the strands weighted?
   e. What were the results of the analysis of each strand?
   f. How were the two analyses integrated? What is the researcher's interpretation?

Guilford Publications
370 Seventh Avenue
New York, NY 10001
212-431-9800
800-365-7006
www.guilford.com



MAP 13.4