Journal of Mixed Methods Research

http://mmr.sagepub.com

A Mixed Methods Investigation of Mixed Methods Sampling Designs in Social and Health Science Research

Kathleen M. T. Collins, Anthony J. Onwuegbuzie and Qun G. Jiao

Journal of Mixed Methods Research 2007; 1; 267
DOI: 10.1177/1558689807299526

Journal of Mixed Methods Research
Volume 1 Number 3
July 2007 267-294

10.1177/1558689807299526
http://jmmr.sagepub.com
hosted at
http://online.sagepub.com

A Mixed Methods Investigation of Mixed Methods Sampling Designs in Social and Health Science Research

Kathleen M. T. Collins
University of Arkansas at Fayetteville
Anthony J. Onwuegbuzie
University of South Florida, Tampa
Qun G. Jiao
Baruch College, City University of New York

A sequential design utilizing identical samples was used to classify mixed methods studies via a two-dimensional model, wherein sampling designs were grouped according to the time orientation of each study's components and the relationship of the qualitative and quantitative samples. A quantitative analysis of 121 studies representing nine fields in the social or health sciences revealed that more studies utilized a sampling design that was concurrent (66.1%) than sequential (33.9%). Also, identical sampling designs were the most prevalent, followed by nested sampling, multilevel sampling, and parallel sampling, respectively. Qualitative analysis suggested that across a number of studies the researchers made statistical generalizations that were not sufficiently warranted—culminating in interpretive inconsistency and contributing to crises of representation, legitimation, integration, and politics.

Keywords: *sampling; sample size; interpretive consistency; representation; generalization*

Since the 1960s, an increasing number of researchers in various fields of social and behavioral sciences have been advocating the combining of quantitative and qualitative approaches to the study of various social phenomena (Creswell & Plano Clark, 2007; Johnson, Onwuegbuzie, & Turner, 2005; Onwuegbuzie & Daniel, 2006; Tashakkori & Teddlie, 1998, 2003a, 2003b). Johnson et al. (2005) referred to this new movement as the *mixing movement* or the *current synthesis stage*. The combining of qualitative and quantitative approaches is most commonly known as mixed methods research. As noted by Johnson and Onwuegbuzie (2004), broadly speaking,

> mixed methods research is formally defined here as *the class of research where the researcher mixes or combines quantitative and qualitative research techniques into a single study*. Philosophically, it is the "third wave" or a third research movement, a movement that moves past the recent paradigm wars by offering a logical and practical alternative. Philosophically, mixed research makes use of the pragmatic method and system of philosophy. Its logic of inquiry

Authors' Note: We are grateful to Dr. Joseph Maxwell, George Mason University, for his constructive feedback on an earlier version of this article, which was presented at the 2006 annual meeting of the American Educational Research Association in San Francisco. Correspondence should be addressed to Kathleen M. T. Collins, Department of Curriculum and Instruction, College of Education and Health Professions, University of Arkansas, 310 Peabody Hall, Fayetteville, AR 72701; phone: (479) 575-4218; e-mail: kcollinsknob@cs.com.

> includes the use of induction (or discovery of patterns), deduction (testing of theories and hypotheses), and abduction (uncovering and relying on the best of a set of explanations for understanding one's results). (pp. 17-18)

The past decade has seen a proliferation in the number of mixed methods studies undertaken. The popularization of mixed methods research has led to an increase in the number of publications devoted to methodological issues in this field—the most comprehensive of which, to date, has been the *Handbook of Mixed Methods in Social & Behavioral Research* (Tashakkori & Teddlie, 2003a), which has provided researchers with some theoretical and practical tools for conducting mixed methods research. Nevertheless, as noted by Teddlie and Tashakkori (2003), six unresolved issues and controversies in the use of mixed methods in the social and behavioral sciences still prevail, namely, (a) the nomenclature and basic definitions used in mixed methods research, (b) the utility of mixed methods research, (c) the paradigmatic underpinning for mixed methods research, (d) design issues in mixed methods research, (e) issues in making inferences in mixed methods research, and (f) the logistics of conducting mixed methods research. Moreover, Onwuegbuzie (2007) has identified the following four crises or challenges that researchers face when undertaking mixed methods research: representation, legitimation, integration, and politics. Each of these challenges, which incorporates one or more of Teddlie and Tashakkori's six unresolved issues, is discussed in the following section.

Four Challenges in Mixed Methods Research

The challenge of representation refers to the fact that sampling problems characterize both quantitative and qualitative research. With respect to quantitative research, the majority of studies utilize sample sizes that are too small to detect statistically significant differences or relationships (i.e., statistical power typically ranging from .4 to .6, with an average statistical power of .5; Cohen, 1962, 1965, 1988, 1992; Schmidt, 1996; Schmidt & Hunter, 1997; Schmidt, Hunter, & Urry, 1976; Sedlmeier & Gigerenzer, 1989) and utilize nonrandom samples that prevent effect-size estimates from being generalized to the underlying population (Onwuegbuzie, Jiao, & Bostick 2004). In qualitative research, the challenge of representation refers to the difficulties researchers encounter in capturing lived experiences via their social texts (Denzin & Lincoln, 2005). In mixed methods research, the challenge of representation often is intensified because both the qualitative and quantitative components of studies bring to the study their own unique challenges. In mixed methods studies, the challenge of representation refers to the difficulty in capturing (i.e., representing) the lived experience using text in general and words and numbers in particular.

The second challenge in mixed methods research pertains to legitimation or validity. With respect to quantitative research, the importance of legitimation, or what is more commonly referred to as "validity," has been long acknowledged and is well documented in the literature, including measurement-related validity (e.g., construct-related validity, criterion-related validity, content-related validity) and design-related validity (e.g., internal validity, external validity; American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999;

Bracht & Glass, 1968; Campbell, 1957; Campbell & Stanley, 1963; Cook & Campbell, 1979; Messick, 1989, 1995; Shadish, Cook, & Campbell, 2001; Smith & Glass, 1987). With respect to the qualitative research paradigm, the issue of legitimation has been more controversial. However, some prominent qualitative researchers argue for "a serious rethinking of such terms as *validity, generalizability,* and *reliability,* terms already retheorized in postpositivist . . . , constructivist-naturalistic, feminist . . . , interpretive . . . , poststructural . . . , and critical . . . discourses" (Denzin & Lincoln, 2005, p. 17) and have reconceptualized traditional validity in ways such as the following: credibility (replacement for quantitative concept of internal validity), transferability (replacement for quantitative concept of external validity), dependability (replacement for quantitative concept of reliability), and confirmability (replacement for quantitative concept of objectivity; Lincoln & Guba, 1985). As is the case for the challenge of representation, the challenge of legitimation is greater in mixed methods studies than in monomethod studies (i.e., quantitative research or qualitative alone). The challenge of legitimation refers to the difficulty in obtaining findings and/or making inferences that are credible, trustworthy, dependable, transferable, and/or confirmable (Onwuegbuzie & Johnson, 2006).

The third challenge in mixed methods research pertains to integration. This challenge compels mixed methods researchers to ask questions such as the following: Is it appropriate to triangulate, expand, compare, or consolidate quantitative data originating from a large, random sample with qualitative data arising from a small, purposive sample? How much weight should researchers and/or consumers place on qualitative data compared to quantitative data?

The fourth challenge in mixed methods research is the challenge of politics. This challenge refers to the tensions that come to the fore as a result of combining qualitative and quantitative approaches. These tensions include any conflicts that occur when different investigators are used for the qualitative and quantitative components of an investigation, as well as the contradictions and paradoxes that come to the fore when qualitative and quantitative data are compared and contrasted. The challenge of politics also pertains to the difficulty in persuading the consumers of mixed methods research, including stakeholders and policy makers, to value the findings stemming from *both* the qualitative and quantitative phases of a study.

Addressing the Four Challenges in Mixed Methods Research

For mixed methods research to maximize its credibility as a paradigm (i.e., investigation validity, communicative validity, and action validity; Kvale, 1995), it is essential that these four challenges be addressed. Of the 13 steps in the mixed methods research process identified by Collins, Onwuegbuzie, and Sutton (2006),[1] we believe that selecting the sampling design, which comprises making decisions about the sampling scheme(s) and sample size(s), is a pivotal step for addressing simultaneously these four challenges. In particular, as noted by Onwuegbuzie and Collins (in press), representation can be improved by ensuring that sampling decisions stem from (a) the research goal (predict; add to the knowledge base; have a personal, social, institutional, and/or organizational impact; measure change; understand complex phenomena; test new ideas; generate new ideas; inform constituencies; or examine the past [Newman, Ridenour, Newman, & DeMarco, 2003]), (b) research objective (i.e., exploration, description, explanation, prediction, or influence [Johnson & Christensen, 2004]), (c) rationale of the study and rationale for mixing qualitative and quantitative approaches (i.e., participant enrichment, instrument fidelity, treatment integrity, and significance enhancement [Collins, Onwuegbuzie, & Sutton, 2006]), (d) purpose of the study and purpose for mixing qualitative and quantitative approaches (e.g., recruit participants, assess the appropriateness and/or utility of existing instrument(s); assess fidelity of intervention; augment interpretation of findings [Collins, Onwuegbuzie, & Sutton, 2006]), and (e) research question(s). For example, with respect to the research goal, the goal of predicting as opposed to understanding complex phenomena likely will lead to a different research objective (i.e., prediction vs. exploration, description, or explanation), rationale (e.g., instrument fidelity vs. participant enrichment), research purpose (e.g., triangulation vs. expansion), and research questions—thereby culminating in different sampling designs, sampling schemes, and sample sizes that yield more rigorous research studies. Representation also can be enhanced by ensuring that the sample selected for each component of the mixed methods study is compatible with the research design.

Furthermore, the samples that are selected for the qualitative and quantitative components should (a) generate adequate data pertaining to the phenomenon of interest under study—these data allow thick, rich description that increases descriptive validity and interpretive validity (Maxwell, 1992); (b) help the researcher to obtain data saturation, theoretical saturation, and/or informational redundancy (Flick, 1998; Lincoln & Guba, 1985; Morse, 1995; Strauss & Corbin, 1990); and (c) allow the researcher to make statistical and/or analytical generalizations. In other words, the sampling design should allow mixed methods researchers to make generalizations to other participants, populations, settings, locations, contexts, events, incidents, activities, experiences, times, and/or processes—such that it facilitates internal and/or external generalizations (Maxwell, 1992). In the context of a mixed methods design, the researcher may purposively extend the conclusions based on the quantitative and qualitative components to the group or context studied (i.e., internal generalizations) or extend the conclusions to another context or group of individuals who are representative of the study's sample (i.e., external or cross-population generalizations) (Hood, 2006; Maxwell, 1992; Onwuegbuzie & Leech, 2005a).

Legitimation can be enhanced by ensuring that inferences stem directly from the underlying sample of units (Curtis, Gesler, Smith, & Washburn, 2000; Kemper, Stringfield, & Teddlie, 2003; Miles & Huberman, 1994; Onwuegbuzie & Collins, in press). In addition, an appropriate sampling design also can increase theoretical validity (Maxwell, 1992). Also, the sampling design can enhance legitimation by incorporating audit trails (Halpern, 1983; Lincoln & Guba, 1985).

The challenge of integration can be reduced by utilizing sampling designs that help researchers to make meta-inferences (i.e., both sets of inferences are combined into a coherent whole; Tashakkori & Teddlie, 2003b) that adequately represent the quantitative and qualitative findings and allow the appropriate emphasis to be placed. Finally, the challenge of politics can be decreased by employing sampling designs that are realistic (i.e., leads to an accurate account of the phenomenon), efficient (i.e., can be undertaken using the available resources), practical (i.e., compatible with the researcher's competencies, experiences, interests, and work style; within the scope of the potential sample members), and ethical (i.e., adheres to the ethical guidelines stipulated by organizations such as institutional review boards to ensure that the integrity of the study is maintained throughout and that all sample members are protected).

Thus, as can be seen, choice of sampling designs is a vital step in the mixed methods research process. Thus, it is somewhat surprising that the issue of sampling was not included as one of Teddlie and Tashakkori's (2003) six issues of concern in mixed methods research. It is even more surprising that a comprehensive review of the literature revealed that, to date, only four articles were identified that deal specifically with the topic of sampling within a mixed methods framework (i.e., Collins, Onwuegbuzie, & Jiao, 2006; Kemper et al., 2003; Onwuegbuzie & Collins, in press; Teddlie & Yu, 2007).

Sampling Designs

As noted by Onwuegbuzie and Collins (in press), sampling designs comprise two major components: the sampling scheme and the sample size. The sampling scheme denotes the explicit strategies used to select units (e.g., people, groups, settings, and events), whereas the sample size indicates the number of units selected for the study. In mixed methods studies, the researcher must make sampling scheme and sample size considerations for both the qualitative and quantitative phases of the study. Thus, mixed methods sampling designs represent the framework within which the sampling occurs, including the number and types of sampling schemes, as well as the sample size.

Sampling schemes. Using the frameworks of Patton (1990) and Miles and Huberman (1994), Onwuegbuzie and Leech (2007) identified 24 sampling schemes that qualitative and quantitative researchers have available for use. All of these sampling schemes fall into one of two classes: random sampling (i.e., probabilistic sampling; $n = 5$) schemes or nonrandom sampling (i.e., nonprobabilistic sampling; $n = 19$) schemes. These 24 sampling schemes are presented in Table 1.[2,3]

Sample size. To increase representation, it is essential that power analyses are conducted in both quantitative (cf. Cohen, 1988) and qualitative (cf. Onwuegbuzie & Leech, 2007) research. Such power analyses provide researchers with information regarding appropriate sample sizes for both quantitative and qualitative phases of a mixed methods investigation. Table 2 presents minimum sample sizes for several of the most common quantitative and qualitative research designs. The sample sizes corresponding to the traditional quantitative research designs (i.e., correlational, causal-comparative, experimental) are the result of the statistical power analyses undertaken by Onwuegbuzie et al. (2004) that represent sizes for detecting moderate effect sizes with .80 statistical power at the 5% level of significance. Conversely, the criteria for sample size in qualitative research are not based on probability computations but represent expert opinion.[4]

Generalizability

According to Onwuegbuzie and Leech (2005a), there are three major types of generalizations: (a) statistical generalizations (i.e., making generalizations or inferences on data

Table 1
Major Sampling Schemes in Mixed Methods Research

Sampling Scheme	Description
Simple[a]	Every individual in the sampling frame (i.e., desired population) has an equal and independent chance of being chosen for the study.
Stratified[a]	Sampling frame is divided into subsections comprising groups that are relatively homogeneous with respect to one or more characteristics and a random sample from each stratum is selected.
Cluster[a]	Selecting intact groups representing clusters of individuals rather than choosing individuals one at a time.
Systematic[a]	Choosing individuals from a list by selecting every kth sampling frame member, where k typifies the population divided by the preferred sample size.
Multistage random[a]	Choosing a sample from the random sampling schemes in multiple stages.
Maximum variation	Choosing settings, groups, and/or individuals to maximize the range of perspectives investigated in the study.
Homogeneous	Choosing settings, groups, and/or individuals based on similar or specific characteristics.
Critical case	Choosing settings, groups, and/or individuals based on specific characteristic(s) because their inclusion provides researcher with compelling insight about a phenomenon of interest.
Theory-based	Choosing settings, groups, and/or individuals because their inclusion helps the researcher to develop a theory.
Confirming/disconfirming	After beginning data collection, the researcher conducts subsequent analyses to verify or contradict initial results.
Snowball/chain	Participants are asked to recruit individuals to join the study.
Extreme case	Selecting outlying cases and conducting comparative analyses.
Typical case	Selecting and analyzing average or normal cases.
Intensity	Choosing settings, groups, and/or individuals because their experiences relative to the phenomena of interest are viewed as intense but not extreme.
Politically important cases	Choosing settings, groups, and/or individuals to be included or excluded based on their political connections to the phenomena of interest.
Random purposeful	Selecting random cases from the sampling frame consisting of a purposefully selected sample.
Stratified purposeful	Sampling frame is divided into strata to obtain relatively homogeneous subgroups and a purposeful sample is selected from each stratum.
Criterion	Choosing settings, groups, and/or individuals because they represent one or more criteria.
Opportunistic	Researcher selects a case based on specific characteristics (i.e., typical, negative, or extreme) to capitalize on developing events occurring during data collection.
Mixed purposeful	Choosing more than one sampling strategy and comparing the results emerging from both samples.
Convenience	Choosing settings, groups, and/or individuals that are conveniently available and willing to participate in the study.
Quota	Researcher identifies desired characteristics and quotas of sample members to be included in the study.
Multistage purposeful random	Choosing settings, groups, and/or individuals representing a sample in two or more stages. The first stage is random selection and the following stages are purposive selection of participants.
Multistage purposeful	Choosing settings, groups, and/or individuals representing a sample in two or more stages in which all stages reflect purposive sampling of participants.

a. Represent random (i.e., probabilistic) sampling schemes. All other schemes are nonrandom (purposive) sampling schemes.

Table 2
Minimum Sample Size Recommendations for Most Common Quantitative and Qualitative Research Designs

Research Design/Method	Minimum Sample Size Suggestion
Research design[a]	
Correlational	64 participants for one-tailed hypotheses; 82 participants for two-tailed hypotheses (Onwuegbuzie, Jiao, & Bostick, 2004)
Causal-comparative	51 participants per group for one-tailed hypotheses; 64 participants for two-tailed hypotheses (Onwuegbuzie et al., 2004)
Experimental	21 participants per group for one-tailed hypotheses (Onwuegbuzie et al., 2004)
Case study	3–5 participants (Creswell, 2002)
Phenomenological	10 interviews (Creswell, 1998); 6 (Morse, 1994)
Grounded theory	15–20 (Creswell, 2002); 20–30 (Creswell, 1998)
Ethnography	1 cultural group (Creswell, 2002); 30–50 interviews (Morse, 1994)
Ethological	100–200 units of observation (Morse, 1994)
Research method	
Focus group	6–9 participants (Krueger, 2000); 6–10 participants (Langford, Schoenfeld, & Izzo, 2002; Morgan, 1997); 6–12 participants (Johnson & Christensen, 2004); 6–12 participants (Bernard, 1995); 8–12 participants (Baumgartner, Strong, & Hensley, 2002)

a. For correlational, causal-comparative, and experimental research designs, the recommended sample sizes represent those needed to detect a medium (using Cohen's [1988] criteria), one-tailed and/or two-tailed statistically significant relationship or difference with .80 power at the 5% level of significance.

extracted from a representative statistical sample to the population from which the sample was drawn), (b) analytic generalizations (i.e., "applied to wider theory on the basis of how selected cases 'fit' with general constructs" [Curtis et al., 2000, p. 1002]), and (c) case-to-case transfer (i.e., making generalizations from one case to another similar case [Firestone, 1993; Kennedy, 1979]). More specifically, statistical generalizability refers to representativeness, whereas analytic generalizability and case-to-case transfer relate to conceptual power (Miles & Huberman, 1994). Sampling designs play a pivotal role in determining the type of generalizations that is justifiable. In particular, whereas large and random samples tend to allow statistical generalizations, small and purposive samples tend to facilitate analytical generalizations and case-to-case transfers. As such, quantitative researchers tend to make statistical generalizations, whereas qualitative researchers tend to make either analytic generalizations or case-to-case transfers. However, in mixed methods research, because quantitative and qualitative research approaches are combined, making appropriate generalizations become even more complex.

Collins, Onwuegbuzie, and Jiao (2006) coined the phrase *interpretive consistency* to denote the consistency between the inferences made by the researcher(s) and the sampling design (e.g., sampling scheme, sample size) used. Figure 1 provides a two-dimensional representation that indicates sampling designs that can yield statistical generalizations with interpretive consistency. In this figure, it can be seen that statistical generalizations typically represent interpretive consistency when both sets of quantitative and qualitative samples are large and, preferably, random—that is, when both sets of samples are representative of the population to which the meta-inferences are being made. When the one

Figure 1
Two-Dimensional Matrix Indicating Sampling Designs That Can Yield Statistical Generalizations That Are Interpretive Consistent

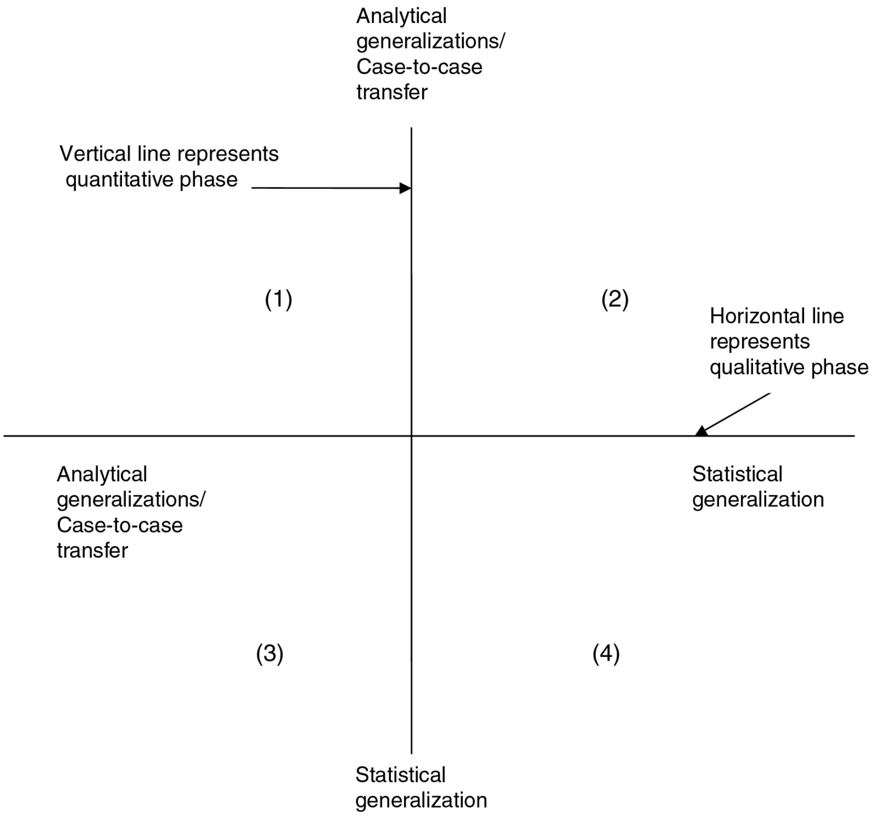
		Qualitative Component(s): Large Sample	Qualitative Component(s): Small Sample
Quantitative Component(s)	Large Sample	Interpretive consistency likely is justified (Type 1)	Interpretive consistency may be justified, especially if the quantitative phase is dominant (Type 2)
	Small Sample	Interpretive consistency may be justified, especially if the qualitative phase is dominant (Type 3)	Interpretive consistency is not justified (Type 4)

Note: The interpretive consistency of statistical generalizations typically is enhanced via the use of random sampling techniques.

set of samples is large (i.e., representative) and the other set small (i.e., purposive), then statistical generalizations might still represent interpretive consistency, especially if the component containing the large sample(s) represents the dominant research approach in the mixed methods study. In contrast, statistical generalizations typically represent interpretive inconsistency when both sets of quantitative and qualitative samples are small. However, for this latter scenario, interpretive consistency likely would occur if analytical generalizations/case-to-case transfers are made instead.

Figure 2 represents a two-dimensional diagram portraying two sets of poles, namely, (a) a vertical pole representing the quantitative phase, with analytical generalization/case-to-case transfer and statistical generalization at the opposite ends of the continuum; and (b) a horizontal line representing the qualitative phase, with analytical generalization/case-to-case transfer and statistical generalization at the opposite ends of the continuum. That is, the vertical pole represents the inferences stemming from the quantitative findings, whereas the horizontal line represents the inferences stemming from the qualitative findings. In Figure 2, the upper left quadrant, labeled as "(1)," represents meta-inferences that involve purely analytical generalization/case-to-case transfer(s). In stark contrast, the bottom right quadrant, labeled as "(4)," represents meta-inferences that involve purely statistical generalizations.

Figure 2
Two-Dimensional Representation of Types of Generalizations Made in Mixed Methods Studies

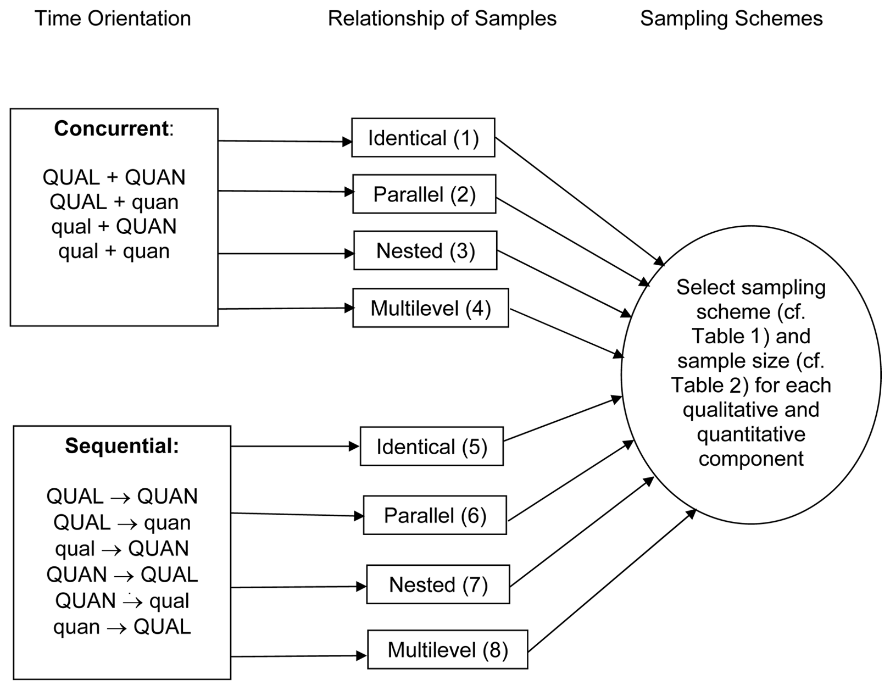


The upper right quadrant, labeled as "(2)," represents meta-inferences that involve a combination of statistical generalizations from the qualitative phase and analytical generalization/case-to-case transfer(s) from the quantitative phase. Finally, the lower left quadrant, marked as "(3)," represents meta-inferences that involve a combination of statistical generalizations from the quantitative phase and analytical generalization/case-to-case transfer(s) from the qualitative phase. Meta-inferences made by mixed methods researchers typically fall into one of these four quadrants, with the exact point on the quadrant being determined by the extent to which statistical generalizations versus analytical generalizations/case-to-case transfers are derived from the data stemming from each phase.

Framework for Mixed Methods Sampling Designs

Onwuegbuzie and Collins (in press) have provided a useful framework for helping mixed methods researchers identify rigorous sampling designs as well as classify mixed

Figure 3
Two-Dimensional Mixed Methods Sampling Model Providing a Typology of Mixed Methods Sampling Designs



Note: QUAL = qualitative; quan = quantitative; + = concurrent; → = sequential. Capital letters denote high priority or weight; lowercase letters denote lower priority or weight.

methods studies in the extant literature with respect to their sampling strategies. This framework incorporates a two-dimensional mixed methods sampling model. Specifically, this model provides a typology in which mixed methods sampling designs can be categorized according to (a) the time orientation of the components and (b) the relationship of the qualitative and quantitative samples. Time orientation refers to whether the qualitative and quantitative phases occur sequentially or concurrently. A sequential sampling design involves the qualitative phase first being conducted to inform the subsequent quantitative phase, or vice versa. That is, a sequential sampling design involves some or all of the findings stemming from the sample selected for one phase being needed to drive the second phase and, subsequently, the sample pertaining to this phase. Conversely, when a concurrent sampling design is utilized, data stemming from the sample selected for one phase do not inform the data stemming from the sample selected for the other phase. However, data obtained from both samples are integrated and interpreted at the data interpretation stage.

According to Onwuegbuzie and Collins (in press), the relationship of the qualitative and quantitative samples either can be identical, parallel, nested, or multilevel. An identical relationship indicates that exactly the same sample members participate in both the qualitative and quantitative phases of the study. In contrast, a parallel relationship denotes that the samples for the qualitative and quantitative components of the investigation are different but are drawn from the same underlying population (e.g., elementary school students from one school in one phase and elementary school students from another school for the other phase). A nested relationship implies that the sample members selected for one component of the inquiry represent a subset of those participants chosen for the other phase of the study. Finally, a multilevel relationship involves the use of two or more sets of samples that are obtained from different levels of the investigation (i.e., different populations). For example, whereas one phase of the study (e.g., quantitative phase) might involve the sampling of students within a high school, the other phase (e.g., qualitative) might involve the sampling of their teachers, guidance counselors, principal, and/or parents.

The two criteria, namely, time orientation and sample relationship, yield eight unique sample designs that typify the various ways that the quantitative and qualitative samples are selected by mixed methods researchers.[5] Figure 3 provides a visual representation of this two-dimensional mixed methods sampling model developed by Onwuegbuzie and Collins (in press). The following section provides a description of study designs that exemplify each of the eight sample designs that comprise the two-dimensional mixed methods sampling model.

Two-Dimensional Mixed Methods Sampling Model

Design 1 involves a concurrent design using identical samples for both quantitative and qualitative components of the investigation. An example of a Design 1 sampling design is the study conducted by Forsbach and Thompson (2003). Utilizing survey methodology, these researchers examined the perceptions of 111 adult survivors of childhood cancer about the quality of their interpersonal adult relationships with peers and siblings. Participants' answers to an online survey were quantitatively analyzed. The survey also contained one open-ended question. Responses to the open-ended question were qualitatively analyzed. By collecting within the same time frame quantitative and qualitative data from the same sample members, the researchers used a concurrent, identical sampling design. Critical case purposive sampling was used to select the identical samples (Miles & Huberman, 1994; Patton, 1990).

Design 2 involves a concurrent design using parallel samples for the qualitative and quantitative components of the study. An example of a Design 2 sampling design is the study conducted by Gordon (2002). This researcher examined music educators' perceptions about teacher stress and the degree that classroom management and discipline are perceived as stressors within the educational environment. Participants comprised two samples of music teachers. The first sample consisted of 103 music educators who completed a questionnaire designed to measure teacher stress levels. The second sample consisted of 4 music teachers who served as case study participants and who provided interview data about job stress. These samples were selected purposively utilizing a homogeneous sampling scheme (Miles & Huberman, 1994; Patton, 1990).

Design 3 involves a concurrent design using nested samples for the quantitative and qualitative facets of the study. An example of a Design 3 sampling design is the study conducted by Mactavish and Schleien (2004). The purpose of this study was to investigate family recreation as a factor impacting dynamics in families that included children with developmental disabilities. In the quantitative phase of the study, 65 families completed a questionnaire. In the qualitative phase of the study, a subset of 16 families who had expressed an interest in participating in a follow-up interview was interviewed. These samples were selected purposively utilizing a criterion sampling scheme (Miles & Huberman, 1994; Patton, 1990).

Design 4 involves a concurrent design using multilevel samples for the quantitative and qualitative components of the study. An example of a Design 4 sampling design is an inquiry undertaken by Schrum, Skeele, and Grant (2002-2003). These researchers utilized a case study approach to evaluate the integration of technology in a college curriculum during a 2-year period from the perspectives of faculty, students, administrators, and technology project directors. In the quantitative component, 13 faculty and students, 183 in Year 1 and 135 in Year 2, completed a prequestionnaire and a postquestionnaire designed to assess their technology skills. In the qualitative component, faculty members, technology project directors, and university administrators participated in focus groups to discuss ways to improve the technology project. The multilevel samples were selected purposively utilizing a criterion sampling scheme (Miles & Huberman, 1994; Patton, 1990).

Design 5 involves a sequential design using identical samples for both the quantitative and qualitative components of the investigation. An example of a Design 5 sampling design is a study conducted by Minor, Onwuegbuzie, Witcher, and James (2002), in which 134 preservice teachers respond to a questionnaire asking them to identify characteristics that exemplify excellent college teaching. They also responded to a second questionnaire designed to assess educational beliefs. In the qualitative phase, preservice teachers' responses pertaining to effective college teaching were developed into themes. In the quantitative stage, preservice teachers' themes and their responses to the educational belief questionnaire were analyzed to determine if a relationship existed between the qualitative themes and preservice teachers' beliefs. The sample was selected purposively utilizing a convenience sampling scheme (Miles & Huberman, 1994; Patton, 1990).

Design 6 involves a sequential design using parallel samples for the quantitative and qualitative components of the study. An example of a Design 6 sampling design is the study conducted by O'Cathain et al. (2004). The study's purpose was to assess the consistency of nurses' triage decisions, the relationship between nurses' clinical backgrounds and decision variations, and nurses' perceptions regarding factors impacting the clinical decision-making process. In the qualitative component, semistructured interviews were undertaken utilizing a sample of 24 nurses to examine, in general, nurses' attitudes toward the decision-support software used in the context of a 24-hour telephone medical helpline, their use of the software to arrive at decisions concerning health care, and the degree that nurses' clinical backgrounds influence the type of advice given to callers. The quantitative component consisted of analysis of software log data collected on all triaged calls during a period of 1 month and the clinical and demographic characteristics of 296 nurses who were employed by the helpline service during this period. Qualitative data were used to identify hypotheses for quantitative component and to provide insight interpreting the quantitative results. These samples were selected purposively utilizing a homogeneous sampling scheme (Miles & Huberman, 1994; Patton, 1990).

Design 7 involves a sequential design using nested samples for the quantitative and qualitative components of the study. An example of a Design 7 sampling design is the inquiry conducted by Igo, Bruning, and McCrudden (2005). These researchers examined the impact of undergraduate students' use of the computer's copy-and-paste function on students' learning outcomes in the context of memory of facts, conceptual learning, and relational inferences. In the quantitative component, students were randomly assigned to two conditions. In the qualitative component, a subset of 24 students was interviewed to determine their thoughts and feeling while engaged in note-taking activities. These samples were selected purposively utilizing a convenience sampling scheme (Miles & Huberman, 1994; Patton, 1990).

Finally, Design 8 involves a sequential design using multilevel samples for the qualitative and quantitative components of the study. An example of a Design 8 sampling design is the study conducted by Drennan (2002). The purpose of the study was to analyze the role of the clinical placement coordinator (CPC) in the context of providing student support in clinical settings. In the qualitative component, nurse education and training centers were randomly selected using a cluster sampling scheme. Individual and focus group interviews were collected from 166 key stakeholders associated with these centers. In the quantitative component, questionnaires were distributed to a sample consisting of 120 CPCs, 300 clinical nurses, and 200 student nurses. This sample was chosen utilizing a random stratified sampling scheme (Miles & Huberman, 1994; Patton, 1990).

Previous Study Examining Mixed Methods Sampling Designs

Collins, Onwuegbuzie, and Jiao (2006) used Onwuegbuzie and Collins's (in press) two-dimensional mixed methods sampling model to investigate the prevalence of sampling designs utilized in mixed methods research. Their second purpose was to use qualitative techniques to examine the consistency between the inferences made by the researcher(s) and the sampling design used (i.e., interpretive consistency).

Collins, Onwuegbuzie, and Jiao (2006) purposively selected (i.e., criterion sampling) all 42 mixed methods studies that were published in the following four leading journals in the school psychology field (i.e., *Journal of School Psychology, Psychology in the Schools, School Psychology Quarterly,* and *School Psychology Review*) over a 4-year period (i.e., 2001 through 2004). Specifically, these researchers selected all mixed methods articles identified by Mihalas, Powell, Onwuegbuzie, Suldo, and Daley (2005), who attempted to determine the prevalence rate of mixed methods articles in the field of school psychology.

Collins, Onwuegbuzie, and Jiao (2006) found that a sequential design using multilevel samples was the most frequent mixed methods sampling design, being used in 40.5% of the studies. More studies involved a mixed methods sampling design that was sequential (66.6%) than concurrent (33.4%). Also, multilevel sampling designs were the most prevalent (54.8%), followed by identical sampling (23.8%), nested sampling (14.3%), and parallel sampling (7.1%), respectively. Their qualitative analysis suggested a degree of interpretive inconsistency in many studies. In particular, regardless of the size of the sample or of the discrepancy between the size of the samples in the quantitative and qualitative phases, the majority of researchers made meta-inferences. Yet in several instances, making such meta-inferences was not sufficiently justified—resulting in interpretive inconsistency.

However, it is not clear how statistically generalizable Collins, Onwuegbuzie, and Jiao's (2006) findings are because they involved mixed methods studies from only one discipline—school psychology. Also, in none of the 42 studies did the researcher(s) explicitly label the study as representing mixed methods research. By not framing their studies in this way, the school psychology researchers likely were not maximizing the extent to which they utilized mixed methodologies. As such, their mixed methods designs were not optimized. Thus, replications of Collins, Onwuegbuzie, and Jiao's study are needed to address the challenges of representation, legitimation, integration, and politics associated with Collins, Onwuegbuzie, and Jiao's results.

Purpose of the Current Study

With this in mind, the purpose of the current mixed methods research was to replicate and extend Collins, Onwuegbuzie, and Jiao's (2006) study. Specifically, in the present inquiry, we investigated the prevalence of sampling designs utilized in mixed methods research. In addition, we examined *interpretive consistency* among these articles. This study represented an extension because it not only involved an examination of a larger sample of mixed methods research articles, but it also involved an examination of studies published in journals representing numerous fields (e.g., psychology, sociology, education, business, nursing). In addition, this investigation utilized studies that were more intentionally mixed methods in nature.

Methods

This study was mixed methods in nature because it utilized both quantitative and qualitative techniques to analyze sampling designs used in mixed methods studies in a multitude of fields. To assess the prevalence rates of the eight mixed methods sampling designs conceptualized by Onwuegbuzie and Collins (in press), we conducted an extensive search utilizing the entire population of electronic bibliographic records of all available fields of social and health sciences at the time of the study. Specifically, 15 electronic bibliographic databases that represent the most widely used electronic sources in the fields of social and health sciences were identified: ABI/Inform Global (ProQuest), Academic Search Premier (EBSCOHost), Business Source Premier (EBSCOHost), CINAHL (EBSCOHost), Education Full-Text (WilsonWeb), ERIC (EBSCOHost), Health Reference Center (Gale InfoTrac), Health Source: Nursing/Academic Edition (EBSCOHost), PsycARTICLES (EBSCOHost), PsycINFO (EBSCOHost), EconLit (EBSCOHost), Education: A SAGE Full-Text Collection (CSA Illumina), Sociological Abstracts (CSA Illumina), Social Services Abstracts (CSA Illumina), and PAIS International (SilverPlatter).

To identify mixed method research articles, the following two keyword search terms were used: *mixed method* and *mixed methodology*. Our rationale for using these keywords was that they would yield articles that were likely to be the most focused with respect to utilizing mixed methods research approaches. Indeed, an examination of the final pool of selected articles supported this prediction because most of these articles (e.g., Daley & Onwuegbuzie, 2004; Hurley, 2001) contained one or more citations of the most prominent methodological works in the field of mixed methods research such as Caracelli and

Greene (1993); Greene, Caracelli, and Graham (1989); and Tashakkori and Teddlie (1998, 2003a).

For each database, we examined all the years for which records existed. Our initial search yielded 703 articles that were then exported into a bibliographic management software (i.e., RefWorks, 2006). After eliminating the 207 articles that appeared in more than one database (i.e., duplicate records), 496 articles remained.[6]

A second screening was conducted based on the list of 496 articles. The abstracts of all these articles were scrutinized for relevance. After eliminating the hidden duplicate records of articles, book reviews, ERIC documents, doctoral dissertations, and pure methodological articles (i.e., nonempirical articles), 128 potentially relevant articles were identified.

Copies of these 128 articles were obtained through either the electronic databases or interlibrary loan services. These articles were then read in their entirety to determine whether they met our inclusion criteria: (a) empirical article that used mixed methods techniques and (b) published in a peer-reviewed English-language journal.[7] Another 7 articles that did not fit these inclusion criteria were eliminated,[8] resulting in a final sample of 121 mixed methods research articles.

We utilized a sequential design using identical samples (Design 5) to conduct our investigation. In this sampling design, the quantitative phase preceded the qualitative phase such that the quantitative phase informed the qualitative phase. For the quantitative phase of our study, the 121 articles were classified using Onwuegbuzie and Collins's (in press) two-dimensional mixed methods sampling model. The frequency rates then were compared and contrasted with respect to (a) the eight mixed methods sampling designs, (b) the time orientation of the components (i.e., concurrent vs. sequential), and (c) the relationship of the qualitative and quantitative samples (i.e., identical vs. parallel vs. nested vs. multilevel). Interrater reliability—a form of *double coding* (Miles & Huberman, 1994)—was used as a means of assessing the validity of the categorizations made. Thus, the verification component of categorization was *empirical* in nature (Constas, 1992). Specifically, two of the researchers independently coded each study by classifying the underlying sampling design into one of the eight mixed methods sampling designs that comprise Onwuegbuzie and Collins's (in press) typology. Kappa's index was used to assess the interrater reliability of the two sets of classifications. This index measures the degree to which observers achieve the possible agreement over and above any agreement than can be expected to occur by chance alone (Cohen, 1960). Because interrater reliability represents a quantitative technique, as well as being empirical, the verification component of categorization was *technical* and was accomplished a posteriori (Constas, 1992). The following criteria were used to interpret the Kappa coefficient: $< .20 =$ poor agreement, $.21\text{-}.40 =$ fair agreement, $.41\text{-}.60 =$ moderate agreement, $.61\text{-}.80 =$ good agreement, $.81\text{-}1.00 =$ very good agreement (Altman, 1991). For our inquiry, Kappa's index was .98 ($SE = .014$), which suggested extremely good agreement between the two raters.

The qualitative phase involved a within-case analysis for each mixed methods study to examine further the time orientation of the quantitative and qualitative components and the relationship between the quantitative and qualitative samples. This analysis involved evaluating the sample size and sampling scheme used for both the qualitative and quantitative components of each mixed methods study. Furthermore, the underlying mixed methods sampling design, sample sizes, and sampling schemes were compared and contrasted with the interpretations, recommendations, conclusions, and meta-inferences made in the discussion section of the article. In particular, an effects matrix (Miles & Huberman, 1994) was used to determine the extent to which the researchers' interpretations, inferences, and/or conclusions were consistent with the overall mixed methods sampling design used (i.e., interpretive consistency). According to Miles and Huberman (1994), effects matrices are useful when there are "ultimate" outcomes and "effects" (p. 137). In the present investigation, the interpretations, conclusions, meta-inferences, and recommendations that were made by the researchers represented the ultimate outcomes. The research question, sample size, sampling scheme, mixed methods sampling design, mixed methods research design, and data analysis techniques primarily represented the effects. As such, this effects matrix allowed us, for each study, to compare directly these effects to the ultimate outcomes. Moreover, this series of matrices facilitated an examination of the extent to which interpretive consistency prevailed among these 121 articles. Excel spreadsheets were used to create these effects matrices, wherein the cell entries in these effect matrices contained either verbatim statements made by the author(s) of the article under examination or our own brief field notes. Consequently, these effects matrices yielded an audit trail (i.e., address the challenge of legitimation), which many qualitative researchers recommended as a method of evaluating legitimation, enhancing legitimation, or both (Halpern, 1983; Lincoln & Guba, 1985; Onwuegbuzie & Leech, 2004). Also, a cross-case analysis (Miles & Huberman, 1994) was undertaken to compare the levels and types of inferences/meta-inferences made across the 121 studies.

According to Onwuegbuzie and Teddlie (2003), there are seven stages of the mixed-methods data analysis process: (a) data reduction, (b) data display, (c) data transformation, (d) data correlation, (e) data consolidation, (f) data comparison, and (g) data integration. Specifically, *data reduction* involves reducing the dimensionality of the qualitative data (e.g., via exploratory thematic analysis, memoing) and quantitative data (e.g., via descriptive statistics, exploratory factor analysis, multiple dimensional scaling, cluster analysis). *Data display*, the second stage, represents describing visually the qualitative data (e.g., matrices, charts, graphs, lists, networks, rubrics, maps, and Venn diagrams) and quantitative data (e.g., tables, graphs). The third stage is the *data transformation* stage, whereby qualitative data are converted into numerical codes that can be analyzed statistically (i.e., *quantitized;* Tashakkori & Teddlie, 1998) and/or quantitative data are transformed into narrative data that can be represented qualitatively (i.e., *qualitized;* Tashakkori & Teddlie, 1998). *Data correlation*, the fourth stage, involves qualitative data being correlated with quantitized data and/or quantitative data being correlated with qualitized data. The fifth stage is *data consolidation*, whereby both qualitative and quantitative data are combined to create new or consolidated variables. The sixth stage, *data comparison*, involves comparing and contrasting data from the qualitative and quantitative data sources. *Data integration*, the seventh and final stage, represents quantitative and qualitative data being integrated into either a coherent whole or two separate sets (i.e., qualitative and quantitative) of coherent wholes. In evaluating the interpretive consistency used in the 121 mixed methods investigations, the data consolidation stage was not used. This was because the quantitative and qualitative data were not combined to develop unique or consolidated variables. However, the remaining six stages of Onwuegbuzie and Teddlie's (2003) model were utilized: data reduction, data display, data transformation, data correlation, data comparison, and data integration.

Table 3
Frequency of Sampling Designs in Mixed Methods Studies Published in Journals From Nine Fields ($N = 121$)

Mixed Methods Sampling Design	Prevalence Rate (%)
Concurrent design using identical samples	28.9
Concurrent design using parallel samples	0.8
Concurrent design using nested samples	19.8
Concurrent design using multilevel samples	16.6
Sequential design using identical samples	7.4
Sequential design using parallel samples	8.3
Sequential design using nested samples	9.1
Sequential design using multilevel samples	9.1

Results

Quantitative Phase

Table 3 provides results from our analysis of the 121 mixed methods studies. It can be seen from this table that by far the most frequent mixed methods sampling design contained in the set of selected investigations was concurrent design using identical samples. With respect to time orientation, more studies utilized a mixed methods sampling design that was concurrent (66.1%) than sequential (33.9%). Furthermore, with regard to the relationship of the qualitative and quantitative samples, identical sampling designs were the most prevalent (36.3%), followed by nested sampling (28.9%), multilevel sampling (25.7%), and parallel sampling (9.1%), respectively.

Table 4 presents a breakdown of the sampling designs across articles representing each of the nine fields. It can be seen from the last column that education was the best represented field. Nursing was the next most represented field, followed closely by social work, public health, psychology, medicine, sociology, business, and library science, respectively. With the exception of the field of business (28.6% vs. 71.4%), the fields were represented by a higher proportion of concurrent sampling designs than sequential sampling designs: education (64.1% vs. 35.9%), library science (100% vs. 0%), nursing (66.7% vs. 33.3%), medicine (88.9% vs. 11.1%), public health (53.9% vs. 46.2%), psychology (72.7% vs. 27.3%), social work (66.7% vs. 33.3%), and sociology (87.5% vs. 12.5%).

Qualitative and Mixed Methods Phases

The majority of the researchers (93.4%) were clear in delineating the sample sizes used in the various phases of their investigations. In the remaining studies, the sample size was not stated clearly in the quantitative phase (2.5%), qualitative phase (5.0%), or both phases (6.6%). Overall, this picture was much more positive than that found for the mixed methods studies from the field of school psychology examined by Collins, Onwuegbuzie, and Jiao (2006). These researchers found that in only 57.8% of the studies were the sample sizes of both the quantitative and qualitative phases specified clearly. Excluding identical samples, which, by definition utilize the same sample for both the quantitative and

Table 4
Mixed Method Sampling Designs Published in Journals From Nine Fields ($N = 121$) (in percentages)

Discipline	CI	CP	CN	CML	SI	SP	SN	SML	Total
Business	0.0 (0)	0.0 (0)	0.8 (1)	0.8 (1)	0.0 (0)	1.7 (2)	0.0 (0)	2.5 (3)	5.7 (7)
Education	7.4 (9)	0.8 (1)	5.8 (7)	6.6 (8)	5.8 (7)	0.8 (1)	4.1 (5)	0.8 (1)	32.2 (39)
Library science	0.8 (1)	0.0 (0)	0.0 (0)	0.0 (0)	0.0 (0)	0.0 (0)	0.0 (0)	0.0 (0)	0.8 (1)
Nursing	5.0 (6)	0.0 (0)	2.5 (3)	2.5 (3)	0.0 (0)	1.7 (2)	0.8 (1)	2.5 (3)	14.9 (18)
Medicine	4.1 (5)	0.0 (0)	0.0 (0)	2.5 (3)	0.0 (0)	0.0 (0)	0.8 (1)	0.0 (0)	7.4 (9)
Public health	2.5 (3)	0.0 (0)	2.5 (3)	0.8 (1)	0.0 (0)	2.5 (3)	1.7 (2)	0.8 (1)	10.7 (13)
Psychology	2.5 (3)	0.0 (0)	3.3 (4)	0.8 (1)	0.0 (0)	0.8 (1)	0.8 (1)	0.8 (1)	9.1 (11)
Social work	4.1 (5)	0.0 (0)	2.5 (3)	1.7 (2)	1.7 (2)	0.8 (1)	0.0 (0)	1.7 (2)	12.4 (15)
Sociology	2.5 (3)	0.0 (0)	2.5 (3)	0.8 (1)	0.0 (0)	0.0 (0)	0.8 (1)	0.0 (0)	6.6 (8)
Total	28.9 (35)	1.7 (2)	19.9 (24)	16.5 (20)	7.4 (9)	7.4 (9)	9.1 (11)	9.1 (11)	100 (121)

Note: I = identical; P = parallel; N = nested; ML = multilevel; C = concurrent; S = sequential. The figures in parentheses represent the number of articles.

Table 5
Mixed Methods Sampling Schemes Published in Journals From Nine Fields ($N = 121$)

	PQUAN	PQUAL	RQUAN	RQUAL	MS	TOTAL
Percentage	90.1	93.4	8.3	5.6	1.7	
Number	109	113	10	7	2	241[a]
Sample size	121	121	121	121	121	121

Note: P = purposive; R = random; QUAN = quantitative; QUAL = qualitative; MS = multistage.
a. Total does not sum to 242 because the quantitative sample for one study was not identified.

qualitative phases, no relationship emerged between the discrepancy of the quantitative and qualitative sample sizes and (a) the time orientation of the components (i.e., concurrent vs. sequential) and (b) whether the sampling design utilized parallel, nested, or multilevel sampling.

With respect to the random sampling scheme, only 14.9% of the studies involved random sampling in one or both phases. In contrast, the vast majority (91.7%) involved some form of purposive sampling in one or both phases of the study.[9] Although due to the lack of sufficient information provided, it was not always clear which of the 19 purposive sampling schemes identified by Onwuegbuzie and Leech (2007) were utilized, it appeared that convenient sampling and criterion sampling represented the most common purposive schemes. Of the 8.3% of studies that utilized random sampling, exactly one half utilized some form of an identical sampling scheme (i.e., either concurrent identical sampling or sequential identical sampling). Table 5 presents the overall prevalence of sampling schemes used (i.e., purposive vs. random) for the quantitative and qualitative phases pertaining to the 121 articles.

Regardless of the size of the sample or of the discrepancy between the size of the samples in the quantitative and qualitative phases, the majority of researchers made meta-inferences (Tashakkori & Teddlie, 2003b), in which both sets of inferences are combined into a coherent whole. Moreover, the majority of these researchers tended to make some type of statistical generalization, rather than making analytic generalizations or case-to-case transfers. Specifically, 58.7% of the articles made some type of statistical generalization—whose justification depended, to a large degree, on the sampling scheme, sample size, and/or sampling design used. The more the statistical generalization made was compatible with these sampling elements, the higher the level of interpretive consistency, and vice versa. For example, an article contained statistical generalizations that were justified (i.e., interpretative consistency) if the set of findings being generalized (i.e., quantitative vs. qualitative phase[s]) stemmed from a large sample. Interpretive consistency was even higher if the sample was random in addition to being large. The level of interpretive consistency across the 121 articles lay on a continuum from interpretive inconsistent to interpretive consistent.

Harwood et al.'s (2003) study provides an excellent example of an article that had high interpretive consistency. These researchers examined how communities, neighborhoods, and stores affect retail pricing and the promotion of beer. Using a multistage, random sampling procedure, the researchers randomly selected 2,024 retail stores from 160 communities. Each store was unobtrusively observed by trained field staff teams who collected information on alcohol process, product placement, extent of related advertising and promotions, and type of size of the store. Using hierarchical regression models and mixed methods data analysis techniques, the researchers found that community, neighborhood, and store characteristics predict beer price, although only community and store characteristics predicted beer promotions. The researchers concluded that "the findings of this study confirm that beer promotion and pricing vary systematically by some characteristics of communities, neighborhoods and stores" (p. 726). Whereas this clearly represents a statistical generalization, involving inferences that go far beyond the underlying sample, it could be argued that the use of a sample that was large and random justified such a generalization—although the researchers might have tempered their findings by using words such as *appears*, *may*, *might*, or *perhaps*—bearing in mind the several limitations that the researchers wisely noted such as the predominance of White or less integrated neighborhoods in the sample. Notwithstanding, the meta-inferences made by the researchers appear to indicate interpretive consistency (cf. Type 1 sampling design in Figure 1 and Quadrant 4 in Figure 2).

Disturbingly, of the 54 studies that included a quantitative and/or a qualitative phase involving a sample of 30 or less participants, 53.7% contained meta-inferences that represented inappropriate statistical generalizations—leading to interpretive inconsistency. Kneebone et al.'s (2003) article provides such an example. These researchers examined the potential of an innovative, scenario-based technique that links a simulated patient with a computer-driven virtual reality (VR) training device for flexible sigmoidoscopy (i.e., visual examination of the lower third of the colon in a search for a nonmalignant growth or tumor). This device provides a safe yet realistic quasi-clinical setting for learners to carry out the procedure while interacting with the "patient." Communication skills are assessed by simulated patients, and quantitative performance data relating to the technique are produced automatically by the VR simulator. Kneebone et al. described their investigation as a pilot study that took place within a nurse practitioner endoscopy course. Seven nurses were selected for the study using a concurrent design with identical samples for both quantitative and qualitative components of the investigation. Qualitative (observation and interviews) and quantitative (communication rating scales and a range of computer-generated output measures from the VR simulator) data were collected. The participants found the procedure to provide a powerful and effective learning experience. However, they all experienced high levels of anxiety. Simulated patients identified strengths in participants' communication skills, alongside areas for development. The simulator-based approach led to an improvement in the quantitative performance measures. Based on the findings pertaining to seven nurses, and despite stating that "we recognise the limitations of our sample size" (p. 57), the researchers made the following unjustified statistical generalization: "Scenario-based training provides a powerful learning experience, allowing participants to build their technical expertise and apply it within a holistic clinical context without the risk of causing harm" (p. 50). This meta-inference, which represented interpretive inconsistency (cf. Type 4 sampling design in Figure 1 and Quadrant 4 in Figure 2), would have been more justified if the authors had used the phrase "scenario-based training *can* provide a powerful learning experience" rather than "scenario-based training *provides* a powerful learning experience." Or more appropriately, the authors could have stated that "the scenario-based training provides a powerful learning experience for the seven nurses." Thus, a subtle change in the phrasing of the meta-inference or the adding of specific information pertaining to the study's sample would have conveyed a more warranted interpretation of the study's findings.

Interestingly, several researchers made interpretive inconsistent meta-inferences despite acknowledging the limitations of their study. For instance, dela Cruz, Brehm, and Harris (2004), who examined focus group tapes and transcripts and three completed quantitative measures (i.e., demographic form, measure of attitudes toward homelessness, and an attendance form) of 15 family nurse practitioners students, acknowledged that "because the sample size was small ($n = 15$), the results of the study can apply only to this sample and cannot be generalized" (p. 553). Despite this statement, the researchers concluded that "a required clinical experience in an HOC [homeless outreach clinic]—as a service-training strategy—significantly changes attitudes toward homeless people and instills in the students social responsibility for the case of one of the fast-growing, underserved, and misunderstood segments of our society" (p. 553).

Discussion

This investigation has replicated and extended the initial study in the area of sampling conducted by Collins, Onwuegbuzie, and Jiao (2006), who used mixed methods techniques to examine the types of sampling designs, sampling schemes, and sample sizes used by mixed methods researchers. That is, as was undertaken in the initial study, this research has involved the use of mixed methods techniques to study the mixed methods literature. However, whereas Collins, Onwuegbuzie, and Jiao focused on studies published in school psychology journals ($N = 42$) in which both qualitative and quantitative techniques were used (typically not explicitly labeled as mixed methods research by the investigators), the current study was much broader in scope, examining studies across nine fields ($N = 121$) that the respective researchers actually labeled as using mixed methods research. Indeed, the 121 studies examined in the present investigation represent the entire *population* of mixed methods studies—at the time of writing.

A sequential design utilizing identical samples was used to classify this population of mixed methods studies on a two-dimensional model in which sampling designs were grouped according to time orientation of study's components and relationship of the qualitative and quantitative samples. A quantitative analysis of 121 studies representing nine fields revealed that a concurrent design using identical samples was the most frequent mixed methods sampling design. Furthermore, twice as many studies utilized a mixed methods sampling design that was concurrent than sequential. Also, identical sampling designs were the most prevalent. These findings are in direct contrast with the results of Collins, Onwuegbuzie, and Jiao (2006), who found that (a) a sequential design using multilevel samples was the most frequent mixed methods sampling design, (b) twice as many studies utilized a mixed methods sampling design that was *sequential* than concurrent, and (c) multilevel sampling designs were the most prevalent. However, these discrepancies likely underscore the difference in the fields examined in the two studies. As noted earlier, whereas in the initial study the researchers analyzed school psychology mixed methods studies, the current investigation cast a much wider net by examining articles from nine fields. The discrepancy in the distribution of sampling designs could reflect the differences in intentionality to conduct mixed methods research, with all the studies examined in the current investigation deliberately being set up as mixed methods investigations. The difference in the distribution of sampling designs also could be indicative of the types of research goals, objectives, rationale of the study, rationale for mixing qualitative and quantitative approaches, purpose of the study, and/or purpose for mixing qualitative and quantitative approaches that distinguish school psychology research from research underlying other fields and disciplines.

The field examined that was closest to the field of school psychology, in terms of discipline or orientation, was psychology. From Table 4, it can be seen that the vast majority (72.7%) of articles classified as belonging to the field of psychology utilized concurrent sampling designs. These articles tended to represent clinical psychology, counseling psychology, and social psychology. It is possible that these branches of psychology (i.e., clinical psychology, counseling psychology) have different rationales for mixing qualitative and quantitative approaches than does the field of school psychology, leading to different emphases on sampling designs. For example, in Collins, Onwuegbuzie, and Jiao's (2006) initial inquiry, the rationale that the researchers used for mixing qualitative and quantitative approaches tended to be either instrument fidelity (i.e., quantitative or qualitative approaches used by the researcher to develop new instruments or maximize the appropriateness and/or utility of new or existing instruments used in the study [Collins, Onwuegbuzie, & Sutton, 2006]) or treatment integrity (i.e., assessing the fidelity of interventions, treatments, or programs; Collins, Onwuegbuzie, & Sutton, 2006). In particular, using qualitative techniques (e.g., interviews, observations) to develop an instrument (i.e., instrument fidelity) or to collect baseline data (i.e., treatment integrity) tends to lead to sequential designs. In contrast, the rationale of a significant proportion of the studies examined in the present investigation was significance enhancement (i.e., enhancing researchers' interpretations of data [Collins, Onwuegbuzie, & Sutton, 2006]). That the majority of studies (54.8%) conducted in the field of school psychology tended to utilize multilevel sampling designs has intuitive appeal because many school psychology researchers tend to sample both students (typically for the quantitative phase) and their school psychologists, teachers, and/or parents (typically for the qualitative stage).

Sampling is even more complex in the mixed methods research process because the meta-inferences that stem from these studies involve combining inferences from the quantitative and qualitative phases of the study that are highly dependent on the sampling scheme and sample size used. Moreover, the quantitative and qualitative components bring into the mixed methods investigation their own challenges of representation, legitimation, integration, and politics. As noted by Onwuegbuzie and Collins (in press), these combined challenges are likely to yield an additive effect or even a multiplicative effect that adversely impacts the quality of meta-inferences made.

In this study, the majority of researchers (i.e., 58.7%) made statistical generalizations. Unfortunately, in many instances (i.e., 53.7% of studies with small samples in at least one phase), such statistical generalizations were not sufficiently warranted—culminating in interpretive inconsistency and contributing to all four crises, namely, representation, legitimation, integration, and politics. The crisis of representation was exacerbated in these studies because the meta-inferences did not adequately represent the findings stemming from the quantitative and/or qualitative phases of these investigations. That is, the meta-inferences made in these inquiries have inadequate external validity (i.e., the extent to which the findings of a study can be generalized across different populations of persons, settings, contexts, and times [Johnson & Christensen, 2004]) and/or external credibility (i.e., confirmability and transferability of findings and conclusions [Maxwell, 1992]). The crisis of legitimation was intensified in these investigations because the meta-inferences did not stem directly from the underlying sample of units (Curtis et al., 2003; Kemper et al., 2003). The crisis of integration was made worse in these inquiries because the meta-inferences did not arise from a Type 1 mixed methods sampling scheme (cf. Figure 1). Finally, the crisis of politics was exacerbated in these studies because the meta-inferences leave these studies open for criticism by the consumers of these studies, including stakeholders and policy makers.

A limitation of this study is that it is not clear the extent to which interpretive consistency may differ as a function of the rationale and purpose guiding the research design within a specific discipline. Subsequently, more research is needed using a larger sample of mixed methods studies to determine why and how different fields and disciplines are mixing qualitative and quantitative approaches. A second limitation of this investigation is that the interpretation of the researchers' meta-inferences may be influenced by the word choice used by the researcher(s) when interpreting the findings. Specifically, semantics involved in the researchers' choice of words when interpreting results may lead to the appearance of overgeneralization even in cases when the researchers' goal was *not* to make generalizations beyond the underlying sample. Unfortunately, it was beyond the scope of the present investigation to determine the researchers' *intentionality* to generalize. At best, we could only ascertain whether the statement made suggested or implied overgeneralization. Future studies should address the extent to which statements that suggest or imply overgeneralization reflect the researchers' intentions.

Notwithstanding, it should be noted that all 42 studies in Collins, Onwuegbuzie, and Jiao's (2006) initial investigation and all 121 studies in the present inquiry could be classified into one of the eight mixed methods sampling designs provides incremental validity to Onwuegbuzie and Collins's (in press) two-dimensional sampling design model—although we recognize that this model might not be applicable for classifying more complex mixed methods sampling designs (cf. Note 5). However, although these two studies have shown the utility of Onwuegbuzie and Collins's model for classifying and understanding the sampling decisions made by researchers in mixed methods studies, the greatest appeal of this model is that it can help researchers to identify a rigorous sampling design for their mixed methods studies, which is a much more complex process than is the case for monomethod studies. In particular, this model can be used to guide sampling decisions made by the researcher, such as selecting a sampling scheme (e.g., random vs. purposive), selecting an appropriate sample size for the quantitative and qualitative phases of the study to enable appropriate generalizations and inferences, and identifying the relationship of the samples (i.e., identical, parallel, nested, multilevel) relative to the study's design. Using our model provides a framework for making these decisions explicit and promotes interpretive consistency between the interpretations made in mixed methods studies and the sampling design used, as well as the other components that characterize the formulation (i.e., goal, objective, purpose, rationale, and research question), planning (i.e., research design), and implementation (i.e., data collection, analysis, legitimation, and interpretation) stages of the mixed methods research process. Another positive aspect of this model is that it does not have a bias toward either quantitative or qualitative approaches, allowing samples to be selected for both phases methodically and rigorously. Thus, we hope that researchers from the social and behavioral sciences and beyond consider using this sampling design model so that they can design their mixed methods studies in a manner that addresses the challenges of representation, legitimation, integration, and politics.

Notes

1. Collins, Onwuegbuzie, and Sutton (2006) conceptualized mixed methods research as comprising the following 13 distinct steps: (a) determining the goal of the study, (b) formulating the research objective(s), (c) determining the research/mixing rationale(s), (d) determining the research/mixing purpose(s), (e) determining the research question(s), (f) selecting the sampling design, (g) selecting the mixed methods research design, (h) collecting the data, (i) analyzing the data, (j) validating/legitimating the data and data interpretations, (k) interpreting the data, (l) writing the final report, and (m) reformulating the research question(s).

2. For a useful alternative typology of sampling schemes, we refer you to Teddlie and Yu (2007), who subdivided sampling schemes into the following four types: probability sampling, purposive sampling, convenience sampling, and mixed methods sampling. In our typology, we included convenience sampling and mixed methods sampling under the label of purposive sampling schemes, which is consistent with many authors (e.g., Hood, 2006; Miles & Huberman, 1994; Onwuegbuzie & Leech, 2007; Patton, 1990).

3. Many authors link sampling schemes with the paradigm. In particular, they associate random sampling schemes with quantitative research designs and purposive schemes with qualitative research designs. However, we think that this dichotomy is problematic for several reasons. First, as noted by Onwuegbuzie and Collins (in press), choice of sampling scheme (i.e., random vs. purposive) should be based on the type of generalization of interest (i.e., statistical vs. analytic). Thus, although not very common, qualitative research can involve random sampling. Similarly, quantitative research can involve purposive sampling—as is the case in the vast majority of quantitative studies (Shaver & Norton, 1980a, 1980b). Thus, we believe that the dichotomy between research paradigm and sampling scheme is false. For a lengthier discussion of this false dichotomy, we refer the readers to Onwuegbuzie and Leech (2005b).

4. It should be noted that the issue of sample size in qualitative research is a controversial one. However, as noted by Sandelowski (1995), a general rule is that sample sizes in qualitative research should not be too small that it is difficult to obtain data saturation, theoretical saturation, or informational redundancy. At the same time, the sample should not be so large that it is difficult to undertake a deep, case-oriented analysis. Teddlie and Yu (2007) referred to this balancing act in qualitative sampling as the representativeness/saturation trade-off.

5. Although Onwuegbuzie and Collins's (in press) typology is comprehensive, representing a large proportion of mixed methods sampling designs, it is not exhaustive. In particular, this typology does not incorporate mixed methods studies in which the quantitative and/or the qualitative phases involve two or more different sampling schemes (e.g., a qualitative phase involving individual interviews and focus groups that necessitate two different sampling schemes and a quantitative phase involving the administration of a survey and an achievement test that necessitate two different sampling schemes). In other words, Onwuegbuzie and Collins's typology does not incorporate mixed methods studies wherein the quantitative and/or the qualitative phases involve mixed purposeful, multistage purposeful, or random multistage purposeful sampling schemes (cf. Table 1).

6. This set of 496 articles revealed that the first study in which the term *mixed methods* was used appeared in 1972 (Parkhurst et al., 1972). However, this study was eliminated in the next round because it represented an unpublished study (i.e., paper presentation).

7. We decided to focus only on published studies that had been peer-reviewed because we believed that they were more likely to have been grounded in mixed methods research due to undergoing a peer-review process.

8. It should be noted that all seven articles that were eliminated at this stage contained the term *mixed method(s)* at least once. However, no evidence was provided that suggested the researcher(s) has used mixed methods approaches. In most cases, these articles had been selected because they had contained the term *mixed method* or its variant in the literature review section.

9. The proportion of studies involving random sampling in one or both phases (i.e., 14.9%) and the proportion of studies involving some form of purposive sampling in one or both phases of the study (i.e., 91.7%) do not sum to 100% because a few studies contained a phase that involved both random sampling and purposive sampling.

References

Altman, D. G. (1991). *Practical statistics for medical research*. London: Chapman and Hall.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing* (Rev. ed.). Washington, DC: American Educational Research Association.

Baumgartner, T. A., Strong, C. H., & Hensley, L. D. (2002). *Conducting and reading research in health and human performance* (3rd ed.). New York: McGraw-Hill.

Bernard, H. R. (1995). *Research methods in anthropology: Qualitative and quantitative approaches*. Walnut Creek, CA: AltaMira.

Bracht, G. H., & Glass, G. V. (1968). The external validity of experiments. *American Educational Research Journal, 5,* 437-474.

Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin, 54,* 297-312.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand-McNally.

Caracelli, V. W., & Greene, J. C. (1993). Data analysis strategies for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 15,* 195-207.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37-46.

Cohen, J. (1962). The statistical power of abnormal social psychological research: A review. *Journal of Abnormal and Social Psychology, 65,* 145-153.

Cohen, J. (1965). Some statistical issues in psychological research. In B. B. Wolman (Ed.), *Handbook of clinical psychology* (pp. 95-121). New York: McGraw-Hill.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Collins, K. M. T., Onwuegbuzie, A. J., & Jiao, Q. G. (2006). Prevalence of mixed methods sampling designs in social science research. *Evaluation and Research in Education*, *19*(2), 83-101.

Collins, K. M. T., Onwuegbuzie, A. J., & Sutton, I. L. (2006). A model incorporating the rationale and purpose for conducting mixed methods research in special education and beyond. *Learning Disabilities: A Contemporary Journal, 4,* 67-100.

Constas, M. A. (1992). Qualitative data analysis as a public event: The documentation of category development procedures. *American Educational Research Journal, 29,* 253-266.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand-McNally.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.

Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson Education.

Creswell, J. W., & Plano Clark, V. L. (2007). *Design and conducting mixed methods research*. Thousand Oaks, CA: Sage.

Curtis, S., Gesler, W., Smith, G., & Washburn, S. (2000). Approaches to sampling and case selection in qualitative research: Examples in the geography of health. *Social Science and Medicine, 50,* 1001-1014.

Daley, C. E., & Onwuegbuzie, A. J. (2004). Attributions toward violence of male juvenile delinquents: A concurrent mixed methods analysis. *Journal of Social Psychology, 144,* 549-570.

dela Cruz, F. A., Brehm, C., & Harris, J. (2004). Transformation in family nurse practitioner students' attitudes toward homeless individuals after participation in a homeless outreach clinic. *Journal of the American Academy of Nurse Practitioners, 16,* 547-554.

Denzin, N. K., & Lincoln, Y. S. (2005). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed., pp. 1-32). Thousand Oaks, CA: Sage.

Drennan, J. (2002). An evaluation of the role of the clinical placement coordinator in student nurse support in the clinical area. *Journal of Advanced Nursing, 40,* 475-483.

Firestone, W. A. (1993). Alternative arguments for generalizing from data, as applied to qualitative research. *Educational Researcher*, *22*(4), 16-23.

Flick, U. (1998). *An introduction to qualitative research: Theory, method and applications*. London: Sage.

Forsbach, T., & Thompson, A. (2003). The impact of childhood cancer on adult survivors' interpersonal relationships. *Child Care in Practice, 9,* 117-128.

Gordon, D. G. (2002). Discipline in the music classroom: One component contributing to teacher stress. *Music Education Research, 4,* 158-165.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11,* 255-274.

Halpern, E. S. (1983). *Auditing naturalistic inquiries: The development and application of a model*. Unpublished doctoral dissertation, Indiana University, Bloomington.

Harwood, E. M., Erickson, D. J., Fabian, L. E. A., Jones-Webb, R., Slater, S., & Chaloupka, F. J. (2003). Effects of communities, neighborhoods and stores on retail pricing and promotion of beer. *Journal of Studies on Alcohol, 64,* 720-726.

Hood, J. C. (2006). The teaching against the text: The case of qualitative methods. *Teaching Sociology, 34,* 207-223.

Hurley, M. M. (2001). Reviewing integrated science and mathematics: The search for evidence and definitions from new perspectives. *School Science and Mathematics, 101,* 259-270.

Igo, L. B., Bruning, R., & McCrudden, M. T. (2005). Exploring differences in students' copy-and-paste decision making and processing: A mixed-methods study. *Journal of Educational Psychology, 97,* 103-116.

Johnson, R. B., & Christensen, L. B. (2004). *Educational research: Quantitative, qualitative, and mixed approaches*. Boston: Allyn & Bacon.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14-26.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2005, April). *Mixed methods research: Is there a criterion of demarcation?* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Kemper, E. A., Stringfield, S., & Teddlie C. (2003). Mixed methods sampling strategies in social science research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 273-296). Thousand Oaks, CA: Sage.

Kennedy, M. (1979). Generalizing from single case studies. *Evaluation Quarterly, 3,* 661-678.

Kneebone, R. L., Nestel, D., Moorthy, K., Taylor, P., Bann, S., Munz, Y., et al. (2003). Learning the skills of flexible sigmoidoscopy—The wider perspective. *Medical Education*, *37*(Suppl. 1), 50-58.

Krueger, R. A. (2000). *Focus groups: A practical guide for applied research* (3rd ed.). Thousand Oaks, CA: Sage.

Kvale, S. (1995). The social construction of validity. *Qualitative Inquiry, 1,* 19-40.

Langford, B. E., Schoenfeld, G., & Izzo, G. (2002). Nominal grouping sessions vs. focus groups. *Qualitative Market Research, 5,* 58-70.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Mactavish, J. B., & Schleien, S. J. (2004). Re-injecting spontaneity and balance in family life: Parents' perspectives on recreation in families that include children with developmental disability. *Journal of Intellectual Disability Research*, *48*(Part 2), 123-141.

Maxwell, J. A. (1992). Understanding and validity in qualitative research. *Harvard Educational Review, 62,* 279-299.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). Old Tappan, NJ: Macmillan.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50,* 741-749.

Mihalas, S., Powell, H., Onwuegbuzie, A. J., Suldo S., & Daley, C. E. (2005). *A call for greater use of mixed methods in school psychology research*. Manuscript submitted for publication.

Miles, M., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

Minor, L. C., Onwuegbuzie, A. J., Witcher, A. E., & James, T. L. (2002). Preservice teachers' educational beliefs and their perceptions of characteristics of effective teachers. *Journal of Educational Research, 96,* 116-127.

Morgan, D. L. (1997). *Focus groups as qualitative research* (2nd ed., Qualitative Research Methods Series 16). Thousand Oaks, CA: Sage.

Morse, J. M. (1994). Designing funded qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 220-235). Thousand Oaks, CA: Sage.

Morse, J. M. (1995). The significance of saturation. *Qualitative Health Research, 5,* 147-149.

Newman, I., Ridenour, C. S., Newman, C., & DeMarco, G. M. P. (2003). A typology of research purposes and its relationship to mixed methods. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 167-188). Thousand Oaks, CA: Sage.

O'Cathain, A., Nicholl, J., Sampson, F., Walters, S., McDonnell, A., & Munro, J. (2004). Do different types of nurses give different triage decisions in NHS Direct? A mixed methods study. *Journal of Health Services Research & Policy*, *9*, 226-233.

Onwuegbuzie, A. J. (2007). Mixed methods research in sociology and beyond. In G. Ritzer (Ed.), *The Blackwell encyclopedia of sociology* (Vol. VI, pp. 2978-2981). Oxford, UK: Blackwell.

Onwuegbuzie, A. J., & Collins, K. M. T. (in press). A typology of mixed methods sampling designs in social science research. *The Qualitative Report.*

Onwuegbuzie, A. J., & Daniel, L. G. (Eds.). (2006). New directions in mixed methods research [Special issue]. *Research in the Schools, 13*(1).

Onwuegbuzie, A. J., Jiao, Q. G., & Bostick, S. L. (2004). *Library anxiety: Theory, research, and applications*. Lanham, MD: Scarecrow Press.

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools*, *13*(1), 48-63.

Onwuegbuzie, A. J., & Leech, N. L. (2004). Enhancing the interpretation of "significant" findings: The role of mixed methods research. *The Qualitative Report*, *9*(4), 770-792. Retrieved November 8, 2005, from http://www.nova.edu/ssss/QR/QR9-4/ onwuegbuzie.pdf

Onwuegbuzie, A. J., & Leech, N. L. (2005a, March). *Generalization practices in qualitative research: Trends in the literature*. Paper presented at the annual meeting of the Eastern Educational Research Association, Sarasota, FL.

Onwuegbuzie, A. J., & Leech, N. L. (2005b). Taking the "Q" out of research: Teaching research methodology courses without the divide between quantitative and qualitative paradigms. *Quality & Quantity: International Journal of Methodology*, *39*, 267-296.

Onwuegbuzie, A. J., & Leech, N. L. (2007). A call for qualitative power analyses: Considerations in qualitative research. *Quality & Quantity: International Journal of Methodology*, *41*(1): 105-121.

Onwuegbuzie, A. J., & Teddlie, C. (2003). A framework for analyzing data in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 351-383). Thousand Oaks, CA: Sage.

Parkhurst, P. E., et al. (1972). *Evaluation of videodisc modules: A mixed methods approach*. East Lansing, MI: National Center for Research on Teacher Learning. (ERIC Document Reproduction Service No. ED 348014)

Patton, M. Q. (1990). *Qualitative research and evaluation methods* (2nd ed.). Newbury Park, CA: Sage.

RefWorks. (2006). RefWorks, release 2006 [Computer software]. San Diego, CA: Author.

Sandelowski, M. (1995). Focus on qualitative methods: Sample sizes in qualitative research. *Research in Nursing & Health*, *18*, 179-183.

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for the training of researchers. *Psychological Methods, 1,* 115-129.

Schmidt, F. L., & Hunter, J. E. (1997). Eight common but false objections to the discontinuation of significance testing in the analysis of research data. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 37-64). Mahwah, NJ: Lawrence Erlbaum.

Schmidt, F. L., Hunter, J. E., & Urry, V. E. (1976). Statistical power in criterion-related validation studies. *Journal of Applied Psychology, 61,* 473-485.

Schrum, L., Skeele, R., & Grant, M. (2002-2003). One college of education's effort to infuse technology: A systematic approach to revisioning teaching and learning. *Journal of Research on Technology in Education, 35,* 256-271.

Sedlmeier, P., & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin, 105,* 309-316.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Shaver, J. P., & Norton, R. S. (1980a). Populations, samples, randomness, and replication in two social studies journals. *Theory and Research in Social Education*, *8*(2), 1-20.

Shaver, J. P., & Norton, R. S. (1980b). Randomness and replication in ten years of the American Educational Research Journal. *Educational Researcher*, *9*(1), 9-15.

Smith, M. L., & Glass, G. V. (1987). *Research and evaluation in education and the social sciences*. Englewood Cliffs, NJ: Prentice Hall.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches* (Applied Social Research Methods Series, Vol. 46). Thousand Oaks, CA: Sage.

Tashakkori, A., & Teddlie, C. (Eds.). (2003a). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.
Tashakkori, A., & Teddlie, C. (2003b). The past and future of mixed methods research: From data triangulation to mixed model designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 671-701). Thousand Oaks, CA: Sage.
Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 3-50). Thousand Oaks, CA: Sage.
Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, *1*, 77-100.