

- Stephenson, W. (1988). Quantum theory of subjectivity. *Integrative Psychiatry*, 6(3), 180-187.
- Stephenson, W., Brown, S. R., & Brenner, D. J. (1972). *Science, psychology, and communication: Essays honoring William Stephenson*. New York: Teachers College Press.
- Stephenson, W., & Burt, C. (1939). Alternative views on correlations between persons. *Psychometrika*, 4, 269-281.
- Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Erlbaum.
- Strauss, A. L., & Corbin, J. M. (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*. Thousand Oaks, CA: Sage.
- Tashakkori, A., & Teddlie, C. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: Sage.
- Thornton, R. K., & Sokoloff, D. R. (1998). Assessing student learning of Newton's laws: The force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula. *American Journal of Physics*, 66(4), 338-352.
- Vantubergen, N. (1975). *QUALANAL user's guide (computer program manual)*. Lexington: Department of Communication, University of Kentucky.
- Waecher, D., Newman, I., & Nolte, D. (1998). Q-factor analysis: The first step in developing typology classifications based upon the Apticom. *Vocational Evaluation and Work Adjustment Journal* (Fall and Winter), 61-66.
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236-244.
- Watts, S., & Stenner, P. (2005). Doing Q methodology: Theory, method and interpretation. *Qualitative Research in Psychology*, 2(1), 67-91.

## 21

## ASSESSING THE QUALITY OF MIXED METHODS RESEARCH

## Toward a Comprehensive Framework

♦ Alicia O'Cathain

**Objectives**

- to describe different conceptualizations of quality of mixed methods research;
- to explore gaps within these conceptualizations and contested areas;
- to construct a comprehensive framework for assessing the quality of mixed methods research;
- to test the quality framework by applying it to a mixed methods study; and
- to identify remaining challenges for assessing quality.

How can one judge whether a mixed methods study has been undertaken well or poorly? It is important to assess the quality of mixed methods research, yet currently, there are no accepted criteria for doing so (Creswell & Plano Clark, 2007). A number of scholars have, however, conceptualized the quality of mixed methods research (Bryman, 2006; Bryman, Becker, & Sempik, 2008; Caracelli & Riggan, 1994; Creswell & Plano Clark, 2007; Dellinger & Leech, 2007; O'Carthain, Murphy, & Nicholl, 2008; Onwuegbuzie & Johnson, 2006; Sale & Brazil, 2004; Tashakkori & Teddlie, 2008; Teddlie & Tashakkori, 2003) and indeed have constructed frameworks for quality assessment (Dellinger & Leech, 2007; Teddlie & Tashakkori, 2009). There is a need now to describe these different conceptualizations of quality and identify any gaps in thinking or contested areas, with the aim of producing a comprehensive framework for assessing the quality of mixed methods research.

#### THE NECESSITY OF A COMPREHENSIVE FRAMEWORK

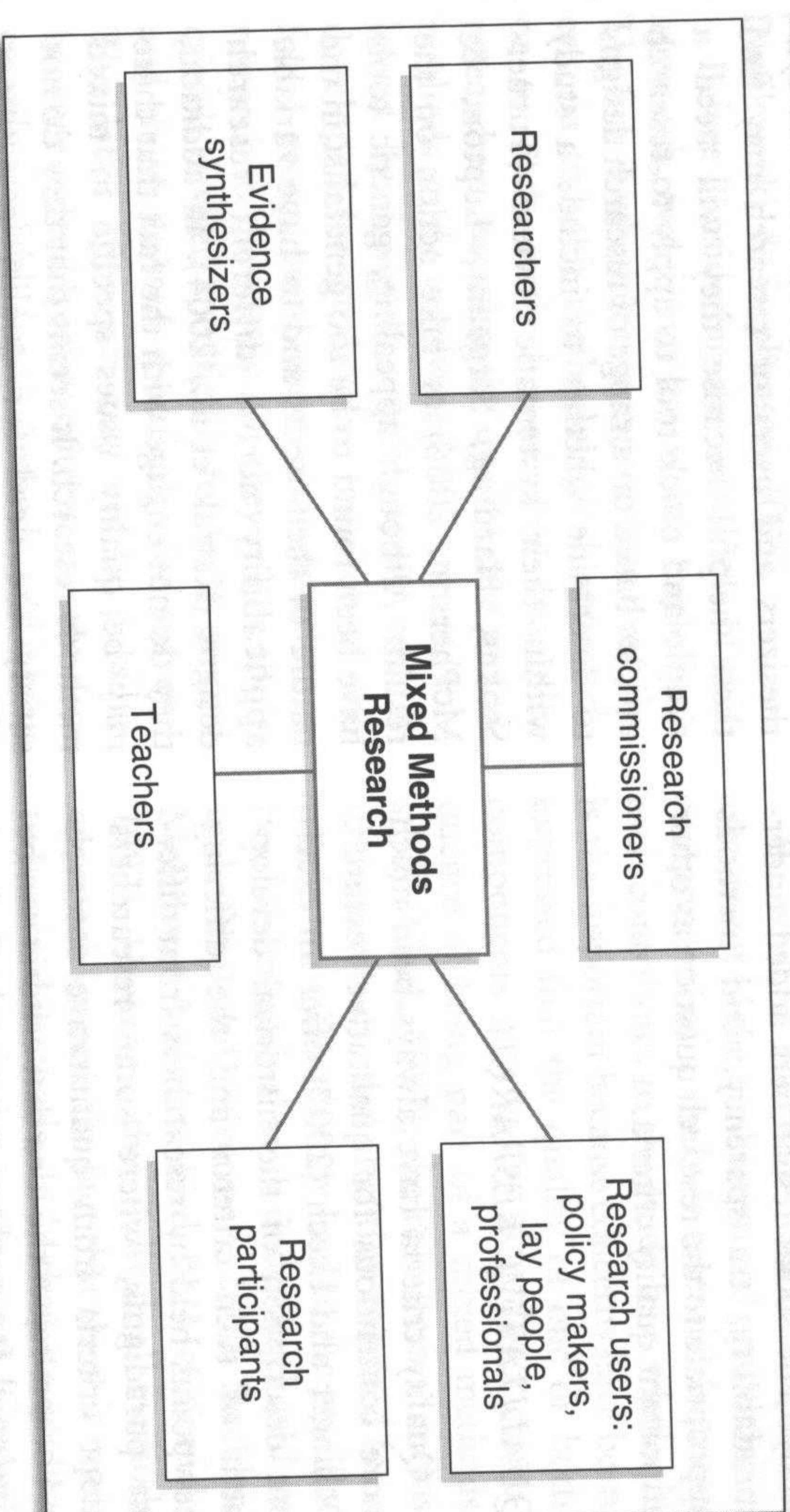
A framework offers a structured description of a complex issue with the purpose of facilitating understanding. In Chapter 17 of this volume Onwuegbuzie and Combs (2010) set out the reasons why an inclusive framework is necessary for data analysis of mixed methods research. These reasons are highly relevant to a framework for quality assessment, including the need to offer guidance to researchers, to establish a common language, and to provide direction for future development.

To be comprehensive, a framework must include the range of discussions on the topic under study. The framework developed within this chapter is based on a critical review of the literature. The search for literature was wide-ranging to ensure that all disciplines undertaking mixed methods research were represented and that expert thinking in

both books and research articles was captured. Constructing the quality framework through a literature review is an approach taken by other researchers addressing the quality of mixed methods research (Pluye, Gagnon, Griffiths, & Johnson-Lafleur, 2009; Sale & Brazil, 2004). Alternative approaches have been taken, in particular conceptualizing quality based on researcher expertise (Teddlie & Tashakkori, 2009) and seeking expert opinion by interviewing individual researchers (Bryman et al., 2008) or undertaking a mapping exercise with a group of researchers (Caracelli & Riggan, 1994). All of these approaches are appropriate and can make complementary contributions to this important topic. With the recent proliferation of methodological publications on mixed methods research, and a growing body of expert researchers, a future priority should be to harness expert opinion through a consensus exercise. Fortunately, this is being addressed through an international Delphi exercise to determine the key quality criteria for mixed methods research (personal communication with Sergi Fabregues Feijoo, assistant professor of qualitative and quantitative research methods, Department of Psychology and Educational Sciences, Open University of Catalonia, Spain).

A comprehensive framework might also be expected to address the needs of the variety of stakeholders who want to assess the quality of mixed methods research (see Figure 21.1). Commissioners of research want to know whether funded studies have addressed the research questions adequately within the allocated resources, that is, delivered the promised goods and gave value for the money. Users of research, such as policymakers, professionals, and lay people, want to know whether they can trust the findings of studies and ultimately take action on them. Research participants want to know whether a study has been a good experience for themselves or others. Teachers of research methods want to communicate core aspects of quality to students of mixed methods research. Researchers want to know

Figure 21.1 Stakeholders Relevant to the Assessment of the Quality of Mixed Methods Research



how to design and execute studies that are methodologically sound and credible. Evidence synthesizers want to use a short instrument to help them to grade the quality of studies for inclusion in systematic reviews. Each of these stakeholders has different needs and is likely to be interested in different aspects of the quality of mixed methods research. Although a comprehensive framework will need to accommodate these perspectives, it is unlikely to offer a solution for all these needs. It is also the case that any framework will be shaped by its author and its potential readership; the framework presented here is researcher-focused with the purpose of helping researchers to undertake and assess good mixed methods research.

qualitative research (Onwuegbuzie & Johnson, 2006; Tashakkori & Teddlie, 2008). This is a useful exercise because it reminds the reader of the established approaches before introducing potential quality criteria for the third methodology (Tashakkori & Teddlie, 2008). In addition, “because researchers collect, analyze, and interpret both forms of data, traditional approaches to validity should not be minimized in mixed methods research” (Creswell & Plano Clark, 2007, p. 146). In this chapter, the discussion of criteria for monomethod studies is deliberately brief to allow more space for discussing mixed methods research. A detailed list of criteria for both qualitative and quantitative research can be found elsewhere (Sale & Brazil, 2004).

#### QUANTITATIVE RESEARCH

##### ♦ Reviewing Approaches to Assessing the Quality of Research

Prior to considering the quality of mixed methods research, authors will often describe the accepted criteria for quantitative and

In 2008, Bryman and colleagues presented a set of “traditional criteria for quantitative research” to social policy researchers in the United Kingdom (Bryman et al., 2008). These criteria were validity, reliability, replicability, and generalizability. Some of these

are relevant to the measures used within a study, some to the data collection and analysis, and others to the inferences from the study. The social researchers added understandability, transparency, and methods appropriate to the research question as other important quality criteria.

### QUALITATIVE RESEARCH

Quality criteria have always been much more contentious for qualitative research. Dellinger and Leech (2007) offer an excellent description of the historical development of such criteria and the differing viewpoints held by researchers from different paradigms, where some researchers adopt criteria from quantitative research and others reject the idea that criteria can be developed for qualitative research. Perhaps the best-known criteria have been developed specifically for qualitative research, addressing the goals of credibility, confirmability, transferability, and dependability (Lincoln & Guba, 1985). Social policy researchers in the United Kingdom added the following to these—transparency, relevance to users, and reflexivity (Bryman et al., 2008).

### MIXED METHODS RESEARCH

Three different approaches can be taken to assessing the quality of a mixed methods study: the generic research approach, the individual components approach, and the mixed methods approach.

#### *The Generic Research Approach*

Does mixed methods research need its own quality criteria? Surely, it is simply a piece of research, and all research can be assessed in the same way. Assessment can be made of a mixed methods study as a whole, using tools developed for generic use across all study designs including monomethod qualitative studies and monomethod quantitative studies. Eleven tools have been found

that purport to be useful for any design (Katrak, Bialocerowski, Massy-Westropp, Kumar, & Grimmer, 2004). Evidence synthesizers and users of research may find these helpful because they will need a simple and quick tool to apply to research articles based on a range of research designs to determine whether to include a study within their systematic review (Turner-Stokes, Harding, Sergeant, Lupton, & McPherson, 2006) or take action on the results. Although appealing, generic tools have been found to be too generalist in the nature of their items and to have variable applicability across different research designs (Katrak et al., 2004). In addition, they do not engage with the fact that there may be quality issues specific to mixed methods research. It seems that they do not satisfy the need for quality assessment among most research stakeholders because only evidence synthesizers have applied this type of assessment to mixed methods studies (Turner-Stokes et al., 2006).

#### *The Individual Components Approach*

Surely, mixed methods research is simply the sum of its qualitative component and its quantitative component. If so, each component can be assessed to ensure it meets the quality criteria appropriate to that methodology. Bryman describes the use of convergent criteria, where the same criteria are used for both components of a study, and separate criteria, where different criteria are used for each of the qualitative and quantitative components (Bryman, 2006). Researchers taking the latter approach have itemized 33 criteria for assessing qualitative methods and 31 for assessing quantitative methods; they envisaged the final set of mixed methods criteria to be a reduced version of these two lists (Sale & Brazil, 2004). A similar approach has been taken in an evidence synthesis study where the qualitative articles were assessed using criteria deemed appropriate to qualitative research, the quantitative articles were assessed using criteria appropriate to quantitative research, and mixed methods

articles were divided into their qualitative and quantitative components and each component assessed separately (Pluye, Grad, Dunikowski, & Stephenson, 2005).

Before considering the usefulness of this approach, two issues are worth exploring. The first is whether it is appropriate to apply criteria to each methodological *approach* or to each *method* used. For example, researchers assessing the quality of a mixed methods study involving focus groups, followed by a survey, might apply criteria developed for qualitative research to the focus group component and quantitative research to the survey. Or they might apply criteria devised specifically for focus groups, and those devised specifically for surveys. This latter approach may not be possible because agreed quality criteria are not necessarily available for all methods, and it may also be challenging if five different methods are employed within a single mixed methods project. However it may be necessary if a method in use is always judged by an agreed set of criteria; a key example of this is the randomized controlled trial in health research (Moher et al., 1995).

The second issue involves the assumption made by some researchers that methods are linked to paradigms—quantitative methods to positivism and qualitative methods to constructivism—and therefore the criteria used to assess different methods should also be linked to paradigms (Sale & Brazil, 2004). Researchers have contested the view that methods are linked to paradigms (Bryman, 1988) and thus that different criteria are needed to assess qualitative and quantitative research (Murphy, Dingwall, Greatbatch, Parker, & Watson, 1998). The same criteria may be relevant, although the appropriate means for judging against these criteria may differ because of the research practices employed in different methodological approaches (Murphy et al., 1998). There has been some empirical exploration of whether to use the same criteria for both qualitative and quantitative research, specifically whether quantitative criteria should be applied to qualitative

research (Bryman et al., 2008): 76% of 226 researchers reported that criteria should be separate and different.

So what is the way forward? Quality assessment of the qualitative and quantitative components of a study is essential because each contributes to the study as a whole. It is also important because concerns have been expressed that the quality of one or both components may suffer as a direct consequence of being part of a mixed methods study (Chen, 1997); the resources in terms of time, money, and attention required for a number of methods may lead to the production of research that is underdeveloped or under analyzed (Silverman, 2000; Stecker, Mcleroy, Goodman, Bird, & McCormick, 1992). However, this individual methods approach ignores the fact that there is more to a mixed methods study than its qualitative and quantitative components (Creswell & Plano Clark, 2007). For example, inferences are drawn from the whole mixed methods study—meta-inferences—not simply from each component (Tashakkori & Teddlie, 2008).

#### *The Mixed Methods Approach*

Attempts have been made to develop quality criteria that address the whole mixed methods study rather than simply the individual components within it—Bryman (2006) calls this the “bespoke” approach where criteria are developed especially for mixed methods studies. The first documented attempt at this focused on mixed methods evaluation. Researchers identified 94 quality criteria, 20 of which were specific to a mixed methods approach (Caracelli & Riggan, 1994). The 20 mixed methods-specific items clustered into four domains—design, data quality and analysis, bias, and interpretation. Examples of items included whether data transformations were defensible, contradictory findings were explained, and convergent findings were not related to shared bias between methods.

Nearly a decade later, two leading scholars in the field proposed what is still the

most comprehensive approach to assessing the quality of mixed methods research (Tashakkori & Teddlie, 2003), and they continued to expand and deepen understanding of their original model (Tashakkori & Teddlie, 2008; Teddlie & Tashakkori, 2009). Tashakkori and Teddlie introduced the concept of *inference quality*, which is a combination of design quality (methodological rigor) and interpretive rigor (authenticity of conclusions from the study). During this period, other researchers produced further conceptualizations of quality, either explicitly building on the Tashakkori and Teddlie model (Dellinger & Leech, 2007; Onwuegbuzie & Johnson, 2006) or undertaking separate endeavors (Creswell & Plano Clark, 2007; O’Cathain et al., 2008).

Onwuegbuzie and Johnson (2006) argued that the Tashakkori and Teddlie model tended to view inference quality as an outcome and that it was essential also to view it as a process, that is, to consider how inferences were drawn as well as the inferences themselves. This led them to add nine types of quality assessment of meta-inferences to the Tashakkori and Teddlie model, including *sample integration legitimation* and *paradigmatic mixing legitimation* (Onwuegbuzie & Johnson, 2006). Teddlie and Tashakkori (2009) did not agree that their conceptualization was limited to quality as an outcome and later argued that they viewed inferences as both processes (steps followed to create meaning) and outcomes (conclusions). Within the same time period, Dellinger and Leech (2007) focused on the meaning of validity in the context of mixed methods research, bringing together Tashakkori and Teddlie’s concept of inference quality and Onwuegbuzie and Johnson’s nine aspects of quality, while adding their own concept of *foundational element*, which they present as a reflection on researchers’ prior understanding of the issue under study. They argue that a researcher’s prior understanding of a phenomenon, gained through reviewing the literature, shapes the research study and the findings and interpretation.

Other scholars have taken a “threats minimization” approach (Creswell & Plano Clark, 2007, pp. 145–149), focused on the extent to which attention is paid to the mixed methods knowledge base (Creswell & Plano Clark, 2007, pp. 162–165), or considered the transparency of reporting (O’Cathain et al., 2008). In the first approach, the researcher must consider potential threats to validity that arise during data collection and analysis (Creswell & Plano Clark, 2007). These threats are design specific and include inadequate data transformation in concurrent designs and use of the same sample sizes for qualitative and quantitative data collection in sequential designs. Creswell and Plano Clark (2007) go on to discuss *mixed methods standards*, where the researcher must draw on the mixed methods knowledge base, for example, showing sensitivity to the challenges of using their mixed methods design. O’Cathain and colleagues (2008) constructed quality criteria for the different aspects of mixed methods studies, namely the design, individual components, integration, and inferences. After applying these to a set of mixed methods studies, they concluded that a lack of transparency while reporting studies hindered quality assessment. They developed a set of criteria for good reporting of a mixed methods study (GRAMMS), based on Creswell’s (2003) earlier guidance for writing a proposal for a mixed methods study.

The conceptualizations of quality from these six groups of researchers have been used explicitly when constructing a quality framework for mixed methods research (Caracelli & Riggin, 1994; Creswell & Plano Clark, 2007; Dellinger & Leech, 2007; O’Cathain et al., 2008; Onwuegbuzie & Johnson, 2006; Tashakkori & Teddlie, 2008). Given the centrality of Tashakkori and Teddlie’s model to the thinking of most of these groups and the comprehensiveness of their approach, it is placed at the core of the framework, and then the contributions of the other five groups are assessed in terms of expanding or challenging this core framework.

Prior to presenting the framework, general issues affecting any framework are explored.

#### ◆ Key Issues to Consider Before Constructing the Framework

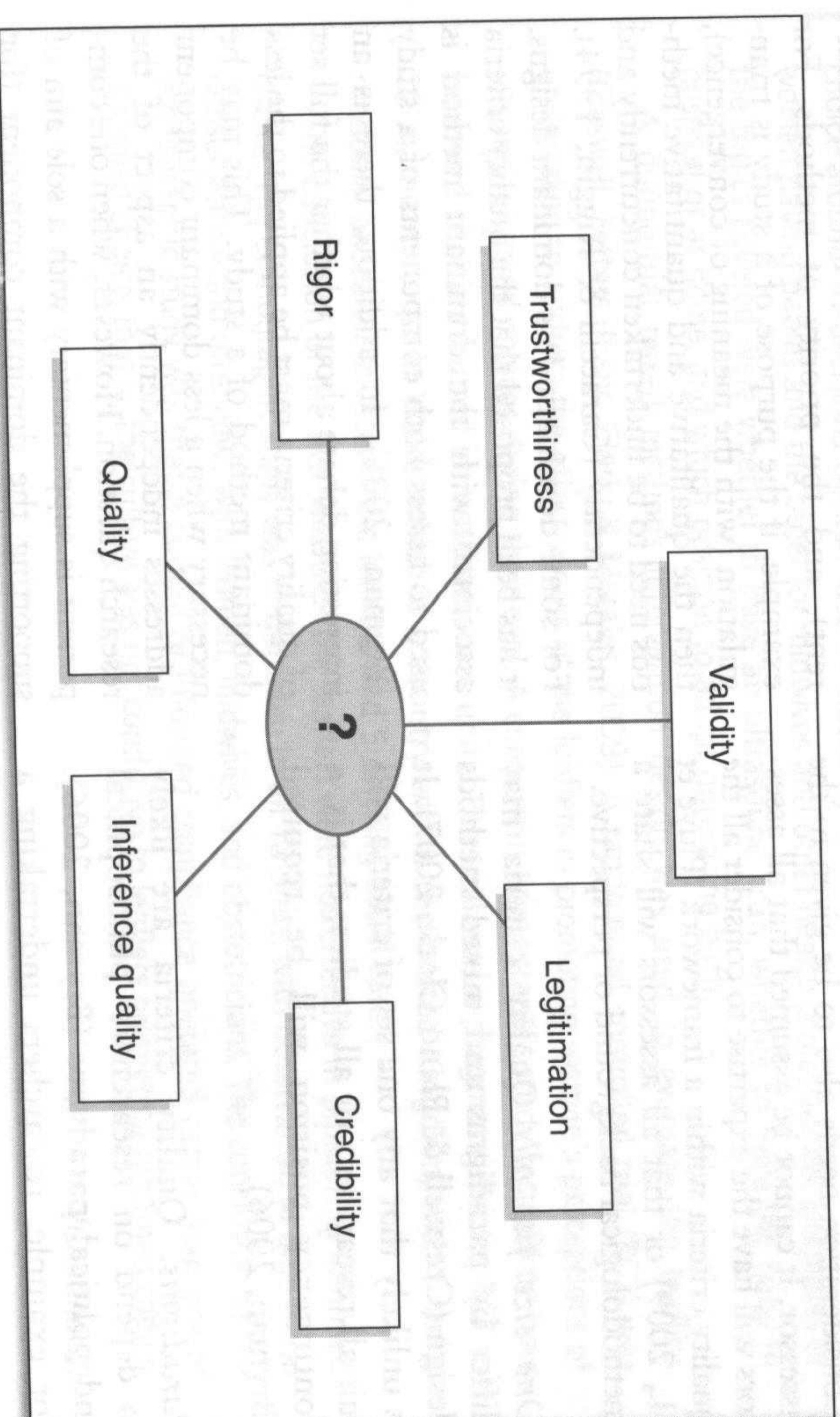
*The language of quality.* Language has been

identified as a challenging aspect of mixed methods research in general (Teddlie & Tashakkori, 2003), and this is highly relevant to the assessment of quality. A language has developed over many years for qualitative and quantitative research, and the quality assessment of both. There is an issue about whether researchers should attempt to use this existing language or create a new language for mixed methods research. The difficulty with using existing language is that it is embedded in the politics of research methods, and therefore, the use of a term associated with one methodology may alienate researchers more aligned with the other methodology. The difficulty

with using a new language is that there is yet more terminology for the poor researcher—and other stakeholders—to understand (Figure 21.2). An excellent solution to this dilemma has been proposed, recommending that researchers introduce new terms only when they have good reason to do so (Tashakkori & Teddlie, 2003, p. 673), for example, when existing terms have been overly used or misused.

Researchers have already introduced a number of terms for the concept of quality in mixed methods research. As Maxwell and Mitterpalli (2010) point out in Chapter 6 of this volume, *validity* has been rejected by some mixed methods scholars, either because it is overused and therefore meaningless (Teddlie & Tashakkori, 2009) or because it is routinely used in quantitative research and therefore disliked by qualitative researchers (Onwuegbuzie & Johnson, 2006). New language of *inference quality* (Teddlie & Tashakkori, 2009) and *legitimation* (Onwuegbuzie & Johnson, 2006) has been introduced to replace validity as an overarching term. There is, however,

Figure 21.2 The Language of Quality



disagreement on this point, with other scholars recommending the continued use of the term *validity* in mixed methods research because it is applied in both qualitative and quantitative research (Creswell & Plano Clark, 2007). There has been further rejection of another overarching term, *rigor*, in favor of *validation* (Giddings & Grant, 2009). The mixed methods community may wish to welcome this diversity, celebrating the variety of language because it reflects a variety of paradigmatic values. However, more consistency of language may facilitate learning and the simple term *quality* might be more helpful for those wishing to assess a mixed methods study as either good or poor.

*The assessed and the assessor.* When quality is considered, there is an assumption that the completed study is being assessed—either the final report or the publications emerging from a study. However, stakeholders also need to assess the quality of research proposals, and a framework needs also to work in this context. Some researchers have explicitly excluded criteria for a proposal when considering quality (Sale & Brazil, 2004), whereas others have described the content of a good mixed methods proposal (Creswell, 2003). Consideration needs also to be given to the assessor. It cannot be assumed that all assessors will have the expertise to consider all the quality criteria within a framework (Pluye et al., 2009) or that all assessors will share a methodological background or perspective.

*One size fits all.* Quality criteria may differ by paradigms and mixed methods design (Creswell & Plano Clark, 2007). It is unlikely that any one set of criteria will suit all researchers or all studies; that is, a contingency position will be required (Bryman, 2006).

*Paradigms.* Quality criteria are likely to depend on researchers' philosophical and political paradigms (Bryman, 2006). For example, researchers undertaking a

transformational mixed methods study would want the community affected to be involved in methodological decisions (Mertens, 2003) and judge a study as poor quality if this did not occur. Giddings and Grant (2009) discuss aspects of quality relevant to mixed methods research undertaken in different paradigms, although not all paradigms common to this approach are included, in particular pragmatism. In Chapter 6 of this volume Maxwell and Mitterpalli (2010) describe the realist perspective on validity and its value to mixed methods researchers. Researchers from some paradigms may not want to engage with the framework presented below or may value some aspects of the framework more than others, with some rejecting the idea that a framework is desirable or possible.

*Design.* Can all mixed methods designs be assessed using the same criteria? A number of researchers have pointed out that criteria may need to be design specific (Bryman, 2006; Sale & Brazil, 2004) and indeed have developed design-specific criteria (Creswell & Plano Clark, 2007). Some criteria will be dependent on the purpose of combining methods (completeness or confirmation), the timing of methods (sequential, concurrent), and the priority of methods. For example, if the purpose of a study is triangulation (with the meaning of convergence), then the qualitative and quantitative methods need to be undertaken concurrently and independently (Caracelli & Riggan, 1994). For some dominant-less dominant designs, it has been proposed that the quality criteria associated with the dominant method is used to assess both components of a study (Bryman, 2006). In addition, there is an interesting debate about whether the full set of quality criteria must be applied to the less dominant method of a study. This may be necessary when a less dominant component addresses independently an aspect of the research question. However, when one component is supplementary with a sole aim of supporting the dominant component (for

example, the role of a qualitative component is to develop a questionnaire only), it may not be appropriate to subject the supplementary component to full quality assessment. This may not be relevant to some definitions of mixed methods research, which exclude this supplementary use of one method.

#### ◆ Proposing a Quality Framework

Having reviewed the approaches taken by researchers when considering the quality of mixed methods research, Tashakkori and Teddlie (2008) appear to offer the most helpful structure for a framework. As stated earlier, their model is placed at the core of

the framework presented in Figure 21.3 (pages 541–544), with the addition of the contributions of other researchers. The concept of inference quality from the Tashakkori and Teddlie model has been replaced with *quality* because although a research team's inferences—and the quality of data on which they are based—must be assessed, methodological rigor should be assessed independently from inferences because if it is good, then stakeholders outside the research team can develop their own inferences and test the interpretative rigor of them. The framework is structured using Caracelli and Riggan's *domains of quality* because they accommodate the range of quality issues discussed by researchers. The domains and items within them are explained below and applied to a mixed methods study described in Box 21.1.

#### BOX 21.1

##### Example Mixed Methods Study: Evaluation of Evidence-Based Leaflets to Promote Informed Choice in Maternity Care

Ten pairs of leaflets were designed to summarize research evidence and promote informed choice around 10 decisions women face in maternity care. One of each pair was written for women having babies and the other for midwives and other health professionals. Examples of the topics covered were whether to have an ultrasound scan, whether to have the baby in hospital or at home, and which position to adopt during the birth.

A during-trial mixed methods intervention design was used to evaluate the leaflets (Creswell, Feters, Plano Clark, & Morales, 2009). A randomized controlled trial (RCT) was undertaken to address the effectiveness of the leaflets in promoting women's perceptions of informed choice during their care. A concurrent ethnographic study focused on how informed choice occurred in maternity care units. The trial was a pragmatic cluster RCT in 10 maternity units in Wales, which is a country within the United Kingdom, with 5 units randomly allocated to receive the leaflets and 5 offering usual care.

The conclusion of the RCT was that the leaflets were not effective in promoting informed choice. The ethnographic study was undertaken in all 10 units, with observation of ante-natal consultations and interviews with women, midwives, and obstetricians. The conclusion of the ethnographic study was that a culture of informed compliance operated rather than one of informed choice, that is, the culture was not conducive to leaflets promoting informed choice. (Continued)

(Continued)

The study was written up in a final report to funders and in a series of journal articles and book chapters. The key articles were a pair of papers published side by side in the same journal, one reporting the RCT (O’Cathain, Walters, Nicholl, Thomas, & Kirkham, 2002) and one reporting the ethnographic study (Stapleton, Kirkham, & Thomas, 2002).

**DOMAIN 1: PLANNING QUALITY**

The first domain addresses how well a mixed methods study has been planned. It can be argued that if attention is not paid to planning the study as a mixed methods design then the study may fail to deliver. This domain is applicable to a research proposal and has four items:

1. The *foundational element* was introduced by Dellinger and Leech (2007). They argue that a comprehensible and critical review of the literature is needed to situate the study, with the research question and study design shaped by the literature review. Caracelli and Riggin (1994) also allude to this when they require the design to be guided by a conceptual framework. Foundational element is also important at the end of a study to situate the findings and indeed may sit appropriately under *interpretive rigor*.

2. *Rationale transparency*—offering a justification for using a mixed methods approach—has been recommended by both Caracelli and Riggin (1994) and Creswell (2003). This is particularly important in the context of strategic use of mixed methods research (Bryman 2007; O’Cathain, Murphy, & Nicholl, 2007), where researchers use mixed methods for the purpose of gaining funding, for example, rather than for its intrinsic value for addressing the research question. This strategic use may lead to neglect of one of the study components, a lack of attention to integration, and no attempt at production of meta-inferences. This item is one of many in the framework related to transparency because researchers

identify this as an important aspect of quality (Bryman et al., 2008).

3. *Planning transparency* is where key aspects of the study including paradigm, design, data collection, analysis, and reporting are detailed in the proposal. Creswell (2003) offers an excellent framework for writing a proposal for a mixed methods study. Many of the issues he recommends for detailed description in research proposals are also relevant to later domains in the framework presented here.

4. A planned study must be *feasible*. Looking at research proposals of mixed methods studies, evidence has been found of large qualitative components planned for execution in short time frames (O’Cathain et al., 2008). Feasibility is not simply an issue for each component of a study but also for the design; it may not be feasible to complete a sequential mixed methods design within a short time frame. Time is not the only resource of importance; there must be enough money, researchers, and expertise available to deliver the study.

Applying these items to the study described in Box 21.1 is difficult because it is a completed study, and this domain is relevant to researchers wishing to write a good proposal and funding agencies wishing to assess whether a proposal is good. However, rationale transparency should also be apparent at the publication stage of a study. Some explanation is given in the article from the qualitative component of the study of why both a quantitative and a qualitative component were necessary: “The effectiveness of these leaflets has been studied in a

Text continued on page 545.

Figure 21.3 Quality Framework for Mixed Methods Research

Stage of Study	Domains of Quality	Items Within Domain	Definition of Item	Source of Domain and Items
Planning	Planning quality	Foundational element	Comprehensible and critical review of the literature is needed to situate the study and shape both the research question and methods.	Dellinger & Leech (2007)
		Rationale transparency	Justification for using a mixed methods approach is provided.	Caracelli & Riggin (1994) Creswell (2003)
		Planning transparency	Details should be given about the paradigm, planned design, data collection, analysis and reporting according to Creswell’s guide for a good proposal.	Creswell (2003)
		Feasibility	The design, and each component, can be undertaken in the resources (time, money, manpower) available.	O’Cathain, Murphy, & Nicholl (2008)
Undertaking	Design quality	Design transparency	Description of design type from known typology, or key aspects of design, if known typologies do not describe design used.	Creswell & Plano Clark (2007) O’Cathain et al. (2008)
		Design suitability	The design is appropriate for addressing the overall research question, matches the reason for combining methods, and is appropriate for the stated paradigm.	Teddle & Tashakkori (2009) Creswell & Plano Clark (2007) Caracelli & Riggin (1994) Onwuegbuzie & Johnson (2006)
		Design strength	The strengths and weaknesses of methods are considered to minimize shared bias and optimize the breadth and depth of the study.	Caracelli & Riggin (1994) Onwuegbuzie & Johnson (2006)
		Design rigor	Methods are implemented in a way that remains true to the design.	Creswell & Plano Clark (2007) Caracelli & Riggin (1994)

(Continued)

Figure 21.3 (Continued)

<i>Stage of Study</i>	<i>Domains of Quality</i>	<i>Items Within Domain</i>	<i>Definition of Item</i>	<i>Source of Domain and Items</i>
Undertaking	Data quality	Data transparency	Each of the methods is described in sufficient detail, including its role within the study.	Creswell & Plano Clark (2007) O'Cathain et al. (2008)
		Data rigor/design fidelity	The extent to which methods are implemented with rigor.	Creswell & Plano Clark (2007) Teddlie & Tashakkori (2009)
		Sampling adequacy	Sampling technique and sample size for each method are adequate in the context of the design.	Creswell & Plano Clark (2007) Onwuegbuzie & Johnson (2006)
		Analytic adequacy	Data analysis techniques are appropriate for the research question and are undertaken properly.	Teddlie & Tashakkori (2009)
		Analytic integration rigor	Any integration taking place at the analysis stage of a study is robust, e.g., data transformations are defensible.	Caracelli & Riggan (1994) Onwuegbuzie & Johnson (2006) O'Cathain et al. (2008) Creswell & Plano Clark (2007)
Interpreting	Interpretive rigor  (Conclusions are based on the findings)	Interpretive transparency	It is clear which findings have emerged from which methods.	O'Cathain et al. (2008)
		Interpretive consistency	Inferences are consistent with the findings on which they are based.	Teddlie & Tashakkori (2009)
		Theoretical consistency	Inferences are consistent with current knowledge or theory.	Teddlie & Tashakkori (2009) Dellinger & Leech (2007)
		Interpretive agreement	Others are likely to reach the same conclusions based on the findings presented, including other researchers and participants.	Teddlie & Tashakkori (2009) Onwuegbuzie & Johnson (2006)

<i>Stage of Study</i>	<i>Domains of Quality</i>	<i>Items Within Domain</i>	<i>Definition of Item</i>	<i>Source of Domain and Items</i>
		Interpretive distinctiveness	Conclusions drawn are more credible than any other conclusions.	Teddlie & Tashakkori (2009)
		Interpretive efficacy	Meta-inferences from the whole study adequately incorporate inferences from the qualitative and quantitative findings and inferences.	Teddlie & Tashakkori (2009) Onwuegbuzie & Johnson (2006) O'Cathain et al. (2008)
		Interpretive bias reduction	Explanations are given for inconsistencies between findings and inferences.	Caracelli & Riggan (1994) Creswell & Plano Clark (2007) Teddlie & Tashakkori (2009)
		Interpretive correspondence	Inferences correspond to the purpose of the study, the overall research question, and the research questions within this.	Teddlie & Tashakkori (2009)
Interpreting	Inference transferability  (Where conclusions can be applied to)	Ecological transferability	Transferability to other contexts and settings.	Tashakkori & Teddlie (2003, 2008, 2009)
		Population transferability	Transferability to other groups and individuals.	
		Temporal transferability	Transferability to the future.	
		Theoretical transferability	Transferability to other methods of measuring behavior.	

(Continued)

Figure 21.3 (Continued)

Stage of Study	Domains of Quality	Items Within Domain	Definition of Item	Source of Domain and Items
Disseminating	Reporting quality	Report availability	Study is successfully completed within allocated resources of time, money, and staff.	Datta (1997)
		Reporting transparency	Key aspects of study reported, according to GRAMMS	Caracelli & Riggin (1994) Creswell & Plano Clark (2007) O'Cathain et al. (2008)
		Yield	Whole more than the sum of the parts.	O'Cathain, Murphy, & Nicholl (2007)
Application in the real world	Synthesizability (Of sufficient quality for inclusion in systematic reviews)	15 quality criteria: 6 for qualitative research 3 for quantitative experimental 3 for quantitative observational 3 for mixed methods	An example criterion is "justification of the mixed methods design."	Pluye, Gagnon, Griffiths, & Johnson-Lafleur (2009)
	Utility	Utility quality	The findings are used by consumers and policy makers.	Caracelli & Riggin (1994) Datta (1997) Dellinger & Leech (2007) Onwuegbuzie & Johnson (2006) Tashakkori & Teddlie (2009)

randomized controlled trial. . . . To understand the social context in which the leaflets were used we undertook qualitative research alongside, but independently of, the randomized trial" (p. 639).

#### DOMAIN 2: DESIGN QUALITY

Design quality is a key component of the Tashakkori and Teddlie model. They propose four criteria within it—design suitability, design fidelity, within-design consistency, and analytic adequacy. In the framework presented in Figure 21.3, design is separated from data collection. Therefore design fidelity and analytic adequacy are moved to Domain 3 on data collection and analysis. Further items, from the work of other researchers, are added to the two remaining Tashakkori and Teddlie criteria:

1. *Design transparency* has been raised as an important aspect of quality in mixed methods studies (Creswell & Plano Clark, 2007; O'Cathain et al., 2008). To offer transparency, a design type may be described from one of the typologies available, or key aspects of the design may be described, given that it is sometimes a challenge to find a design type that fully describes a study in practice. Aspects of design that should be described are priority of approaches, purpose of combining methods, sequencing of methods, and stage at which integration takes place. Creswell and Plano Clark (2007) recommend a visual diagram of the design to facilitate transparency.

2. *Design suitability* or appropriateness has been put forward within the Tashakkori and Teddlie model and by a number of other scholars. The design must be appropriate for addressing the overall research question. Each method must also be appropriate for addressing research questions within the overarching research question. The design must match the stated purpose for combining methods, as well as the research question (Caracelli & Riggin, 1994; Teddlie

& Tashakkori, 2009). Finally, the design must fit with any stated paradigm—called *paradigmatic mixing legitimacy* (Onwuegbuzie & Johnson, 2006).

3. *Design strength* is based mainly on items from Caracelli and Riggin (1994), with some input from Onwuegbuzie and Johnson (2006). They are concerned that researchers design studies that optimize breadth (associated with quantitative research) and depth (associated with qualitative research); consciously consider how the weakness of one method is compensated by the strengths of the other (called *weakness minimization legitimacy* by Onwuegbuzie & Johnson, 2006); and select methods to minimize shared bias. This may be design specific, for example, that methods used for different but complementary purposes enable a more comprehensive study (Caracelli & Riggin, 1994).

4. *Design rigor* is where the methods are implemented in a way that remains true to the design (Creswell & Plano Clark, 2007). For example, rigor is compromised in triangulation designs if methods are not implemented concurrently and independently (Caracelli & Riggin, 1994). Creswell and Plano Clark (2007) offer design-specific recommendations for attaining rigor in concurrent and sequential designs.

Applying this domain to the example in Box 21.1, the design is not transparent within the journal articles, and the study fails to meet this quality item. The design was a randomized controlled trial and ethnographic study undertaken concurrently with the purpose of complementarity. Within a typology of mixed methods intervention studies, it is now known as a "during-trial" design (Creswell, Fetters, Plano Clark, & Morales, 2009). This design was appropriate for addressing the overall research question of understanding the effectiveness of evidenced-based leaflets in maternity care. However, this overall research question is never stated within the journal articles from the study. The individual questions addressed by the



qualitative and quantitative methods are stated, and each method is suitable for addressing its question (see planning quality). The methods fit together, sitting side by side as they address separate but interrelated research questions. The design strength is very good—the trial offers measurement of the effectiveness of the leaflets across five maternity units, and the ethnographic study offers depth of understanding of the culture in which the leaflets were used. The approach to minimization of bias is not obvious for this design. Paradigms are never discussed and therefore paradigmatic legitimation cannot be assessed. The design required that two methods were implemented to a high quality, and this occurred, but generally, design rigor is questionable due to the lack of transparency about the design and therefore the explicit attention paid to design rigor. In particular, did the design require that data sharing occur between qualitative and quantitative researchers throughout the study or only at the interpretation stage? In conclusion, the study is assessed poorly on this domain due to a lack of attention paid generally to the mixed methods design.

### DOMAIN 3: DATA QUALITY

The domain of data quality includes data collection and analysis and has five items:

1. *Data transparency* is where each of the methods is described in detail, including its role within the study, data collection, sampling, sample size, and analysis (Creswell & Plano Clark, 2007; O'Cathain et al., 2008).

2. *Data rigor—or design fidelity*—concerns the extent to which methods are implemented with rigor (Teddlie & Tashakkori, 2009). Here it is important to consider whether a method has been compromised because it is part of a mixed methods study (O'Cathain et al., 2008). It may not be as developed as it needs to be due to lack of resources. For example, a

Delphi technique only has two rounds because it was part of a sequential mixed methods study, and there was not time for the three rounds considered appropriate for the research question. Creswell and Plano Clark (2007) identify potential threats to validity at the data collection stage of two key mixed methods designs.

3. *Sampling adequacy* is where the sampling technique and sample size are adequate for each method in the context of the design (Creswell & Plano Clark, 2007). This is extremely important for the later domains of interpretive rigor and inference transferability because *sample integration legitimation* impacts on the quality of any meta-inferences (Onwuegbuzie & Johnson, 2006). Researchers may find themselves generalizing their findings inappropriately because they have not paid attention to the type of sample and sample size required for each method (e.g., large random sample for quantitative component) or the relationship between the qualitative and quantitative samples.

4. *Analytic adequacy* means that data analysis techniques are appropriate for the research question and undertaken properly (Teddlie & Tashakkori, 2009). For example, the right statistical tests have been used for the quantitative component.

5. *Analytic integration rigor* applies to the quality of any integration taking place at the analysis stage of a study. This might involve data transformation of qualitative data to quantitative data, or more rarely, quantitative data to qualitative data. It might also involve the use of findings from one component of a study to guide the analysis of another component, or placing both types of data in a matrix for within-case and across-case analysis. It is really a part of analytic adequacy but is presented separately here because so many scholars have identified it as a challenge specific to mixed methods research. Data conversion quality in particular has concerned a number of scholars (Caracelli & Riggan,

1994; Creswell & Plano Clark, 2007), including its impact on meta-inferences (Onwuegbuzie & Johnson, 2006). Helpful guidance is emerging on quantifying, the most common approach to data conversion (Sandelowski, Voils, & Knafl, 2009).

Applying this domain to the example in Box 21.1, data transparency, data rigor, sampling adequacy, and analytic adequacy all appear to be excellent for each method. There is no integration at the analysis stage, and so the analytic integration rigor item is not relevant. The lack of any integration between components, including inferences from both the qualitative and quantitative findings, means that the effect of sampling adequacy on meta-inferences will be irrelevant for the next two domains.

### DOMAIN 4: INTERPRETIVE RIGOR

The quality of inferences is very important to users of research, who must find them credible and trustworthy if they are to take action on them (Tashakkori & Teddlie, 1998). Researchers have considered the complexity of this issue (Miller, 2003), and a call has been made for standards for the evaluation of the accuracy or authenticity of conclusions from mixed methods studies (Tashakkori & Teddlie, 2003). Interpretive rigor considers whether conclusions are based on the findings of the study, with the following eight items:

1. *Interpretive transparency* is where it is clear which findings have emerged from which methods (O'Cathain et al., 2008). Without this, links cannot be made between data quality and inferences.

2. *Interpretive consistency* concerns whether inferences are consistent with the findings on which they are based (Teddlie & Tashakkori, 2009). In addition, a number of inferences may be drawn from a small set of findings, and these inferences must be consistent with each other.

3. *Theoretical consistency* is where the inferences are consistent with current knowledge or theory (Teddlie & Tashakkori, 2009). Dellinger and Leech (2007) acknowledge that their *inferential consistency*, where inferences are consistent with what was already known, is very similar to this.

4. *Interpretive agreement* means that others are likely to reach the same conclusions based on the findings presented, including other researchers and study participants (Teddlie & Tashakkori, 2009). The process by which this is attained may be inside-outside legitimation, where meta-inferences are considered by peer review of an outside party and member checking for an insider view so that inferences do not rely only on the research team (Onwuegbuzie & Johnson, 2006).

5. *Interpretive distinctiveness* considers whether the conclusions drawn are more credible than any other conclusions (Teddlie & Tashakkori, 2009). A researcher must be able to discount other possible interpretations. Strategies for achieving this for the individual components within a mixed methods study include negative case analysis in the qualitative research and controlling for variables in the quantitative research.

6. *Interpretive efficacy* is where the meta-inferences from the whole study adequately incorporate inferences from the qualitative and quantitative findings and inferences (Teddlie & Tashakkori, 2009). Other researchers have given consideration to the balance of inferences from different components of a study (Caracelli & Riggan, 1994; O'Cathain et al., 2008; Onwuegbuzie & Johnson 2006). Onwuegbuzie and Johnson (2006) call an aspect of this *political legitimation* and describe a violation of this when different researchers undertake the qualitative and quantitative components, which then affects the conclusions drawn if one is more powerful or likely to interpret contradictions in data and findings in a particular way. This draws attention to who does the integration (O'Cathain et al., 2008) and also

the extent to which processes of integration are visible within journal articles emerging from a study. Another aspect of what could also be called inference balance is sample integration legitimation, where attention is paid to the way in which individuals are sampled for each component when making meta-inferences (Onwuegbuzie & Johnson, 2006).

7. *Interpretive bias reduction* is a criterion within Tashakkori and Teddlie's interpretive efficacy. It has been drawn out here because a number of researchers have discussed this as an important aspect of quality (Erzberger & Kelle, 2003). Teddlie and Tashakkori (2009) request that explanations are given for inconsistencies between inferences. Caracelli and Riggan (1994) recommend that interpretation of the data collected by different methods should consider bias of the methods, in particular that nonconvergent findings are plausibly explained and that convergent findings are not the result of shared bias between the methods. Further exploration of contradictory findings may be particularly important for concurrent designs (Creswell & Plano Clark, 2007).

8. *Interpretive correspondence* means that inferences correspond to the purpose of the study, the overall research question, and the research questions within this (Teddlie & Tashakkori, 2009). The extent to which the researchers have answered their research question and met any other goals of the research must be assessed.

Applying these to the example in Box 21.1, the study performed well on some of the items within this domain. The interpretive transparency was excellent due to the separate reporting of the qualitative and quantitative components of the study; this made it obvious which findings were related to which methods. There was also interpretive consistency in that inferences were consistent with the findings on which they were based. For the trial, the inference was that the leaflets were not

#### DOMAIN 5: INFERENCE TRANSFERABILITY

Tashakkori and Teddlie have proposed the extremely useful concept of inference transferability for mixed methods research—the degree to which the conclusions can be applied to other entities or settings. This is equivalent to *external validity* for quantitative research and *transferability* for qualitative research (Teddlie & Tashakkori, 2003). Inferences can be drawn from each component of a study, with external validity/generalizability considered for the quantitative component and transferability considered for the qualitative component. Mixed methods studies also have meta-inferences, which are the inferences from the whole study rather than simply the individual components. Teddlie and Tashakkori (2009) propose four types of transferability: ecological (transferability to other contexts and settings), population (transferability to other groups and individuals), temporal (transferability to the future), and theoretical (transferability to other methods of measuring behavior).

Taking the example in Box 21.1, inference transferability cannot be considered using the journal articles from the study because the qualitative and quantitative components were published separately. The journal article from the quantitative component paid little attention to generalizability. However, there was transparency of sampling and description of the sample to allow the reader to consider generalizability. The main issue was the sampling of maternity units—there was a mixture of small and large units in a single country, Wales. The study was generalizable to different sizes of maternity units but not necessarily to other countries in the United Kingdom and the rest of the world, unless maternity care operated in a similar way in these countries. Transferability was not considered explicitly in the qualitative article, although there was transparency of sampling and description of participants to allow the reader to draw conclusions about

this. The qualitative component identified the importance of culture within maternity units, which was not conducive to promoting the outcome important to the leaflets. The inference from the quantitative component was that leaflets were not effective in promoting informed choice. The inference from the qualitative component was that the maternity units did not operate a culture of informed choice but rather one of informed compliance. The meta-inference—never explicitly stated in the journal articles emerging from the study—was that leaflets were not effective in maternity units with a culture of informed compliance. The qualitative component offered important information to allow research users to consider whether leaflets might work for them. If they considered that a culture of informed compliance rather than informed choice was in operation in their country, or maternity unit, then they could transfer the findings of the RCT to their context and not purchase leaflets for their service users.

#### DOMAIN 6: REPORTING QUALITY

1. *Report availability* is a factor in that those who commission research will judge quality by whether a study has been successfully completed, and this has occurred within the allocated resources of time, money, and staff (Datta, 1997). This item is relevant to all types of studies but may be more important to ask of mixed methods research because these studies may be more complex and more expensive than other types. Datta describes a failed mixed methods study where the final comprehensive report was delayed for many years, the cost overrun was high, there was staff burnout, expectations were not met, and only some parts of the study were reported (Datta, 1997).

2. *Reporting transparency* means that key aspects of the study are clearly and explicitly reported. If they are not, then

assessment cannot be made of the above quality domains. Researchers writing up their studies may wish to follow guidelines on good reporting of a mixed methods study (GRAMMS) (O'Cathain et al., 2008). Creswell and Plano Clark (2007) promote the need for attention to key aspects of the mixed methods knowledge base, and underlying this is the need for transparency. Caracelli and Riggan (1994) also identified the need to report findings in a way that maximizes the interest of stakeholders.

3. *Yield* refers to the knowledge gained from a mixed methods study over and above the knowledge gained from undertaking two independent qualitative and quantitative studies (O'Cathain et al., 2007). This may not occur within a study because researchers fail to integrate different components of a study or to make what is learned from integration explicit within their report of a study.

For the example study in Box 21.1, the report was available within a few months of the study ending. The report itself was published by an external body, which increased its availability, and a number of journal articles emerged from the study. Although reporting of each component was transparent within journal articles, there was no transparency of the mixed methods aspects of the study such as design and integration. Therefore, the yield of the study was the sum of its parts. A simple integration of findings to produce a meta-inference from the study would have been of considerable benefit.

#### DOMAIN 7: SYNTHESIZABILITY

In health research, there is a tradition of synthesizing evidence on the effectiveness of drugs and other treatments. This usually involves systematically searching for all randomized controlled trials of the treatment under study, undertaking a quality assessment of each trial by scoring quality using a validated set of criteria, and either excluding

studies of low quality prior to a meta-analysis of remaining studies or ranking studies by quality within the synthesis. Quality assessment is a key part of this process. A number of methodological approaches to synthesizing qualitative studies have emerged (Paterson, Thorne, Canam, & Jilings, 2001). Here, the issue of quality assessment is contested, with concerns that checklists cannot be applied to the diversity of methods within qualitative research and arguments about whether poor quality studies should or should not be excluded from reviews. There is also a recognition of a need to synthesize evidence from a range of study types including those based on qualitative, quantitative, and mixed methods research (Dixon-Woods, Agarwal, Jones, Young, & Sutton, 2005; Pope, Mays, & Popay, 2007). Harden and Thomas (2010) explore this in depth in Chapter 29 of this volume. The term *mixed studies review* has been introduced for this type of synthesis (Pluye et al., 2009).

Researchers need to determine whether a mixed methods study is worth including in an evidence synthesis or what weight should be given to it within the synthesis. Within mixed studies reviews, there is a need to assess articles reporting only qualitative research, only quantitative research, and combinations of both. Pluye and colleagues critically examine the quality appraisal tools that have actually been applied in mixed studies reviews, finding 12 formal quality appraisal procedures used in 17 systematic mixed studies reviews in the health sciences, although no validated checklists were found. From this, they propose a set of 15 quality criteria, with a scoring system. Their aim is a minimum set of criteria for ease of use, rather than an exhaustive list. The criteria include six for application to qualitative studies or the qualitative components of mixed methods studies, six for application to different types of quantitative studies, and three for mixed methods studies for use in conjunction with the qualitative and quantitative criteria.

If the instrument constructed by Pluye and colleagues was applied to the example study in Box 21.1, then the qualitative component and quantitative component would be assessed separately because they were published separately. That is, the three criteria relevant to mixed methods would not be used. When the six criteria for qualitative research are applied to the ethnographic study, five are met; the exception is "discussion of researchers' reflexivity." The qualitative researchers certainly practiced reflexivity, but the word count permitted by the journal was so small that a decision was made not to discuss reflexivity because it was not necessary in the context of the value set of the journal. Two of the three criteria relevant to quantitative experimental research are met when applied to the RCT, with the exception of blinding; blinding was not appropriate because this was a pragmatic trial. Thus, the study scores 5 out of 6 (83%) for the qualitative component and 2 out of 2 (100%) for the quantitative component. If it were assessed as a mixed methods study, it would score 7 out of 8 for its components and 0 out of 3 for the mixed methods aspects, totaling 7 out of 11 (64%).

#### DOMAIN 8: UTILITY

A number of researchers have put forward the utility of a study as an indicator of quality. Datta (1997) considers whether the results are usable, Caracelli and Riggan (1994) whether the combination of methods informs changes in policy, Onwuegbuzie and Johnson (2006) whether consumers and policy makers use the meta-inferences—called political legitimization—and Dellinger and Leech (2007) whether historically the results are used—called the historical element. Of course, poor research can be used by policy makers, and indeed, Dellinger and Leech propose caution for this reason. A related issue is what Dellinger and Leech (2007) call the consequential element,

which is the social acceptability of the consequences of using findings from a study. An example might be a finding that breast care nurses are not effective in helping young women deal with postoperative care. If policy makers withdraw funding for the service, then this might be seen as unacceptable by some charities and subgroups of the population.

Consideration of the utility of a study is difficult in practice. It may be challenging to associate specific actions by research users with specific studies. Some studies may have an immediate impact because they are newsworthy and therefore disseminated widely. Other studies may contribute quietly to a growing evidence base about a particular issue. In the case of the example of the leaflets in maternity care, a study would have to be made of the continuing use of leaflets and changes in the culture of maternity care, perhaps by considering policy documents in this area or surveying maternity units. Mixed methods scholars have proposed developing a utilization quality audit for the evaluation of utilization quality (Tashakkori & Teddlie, 2009).

#### ITEMS NOT INCLUDED IN THE FRAMEWORK

Some quality criteria identified have not been included in the framework. For example, Onwuegbuzie and Johnson (2006) propose *sequential legitimization*, where inferences drawn may depend on the sequencing of methods. They suggest that sequencing could be reversed to test this. However, sequencing is chosen to best address the research question—that is, the research question might call for parallel designs—so this solution is limited and therefore is not included in the framework. O'Cathain et al. (2008) discuss the need for expertise in the individual methods within a study. However, the weakness or strength of data collection will be obvious

from the reports or journal articles being assessed regardless of the level of expertise present on any research team.

### ◆ Challenges

**Too many criteria.** The experience of trying to apply all of the items within this comprehensive framework to a real-life mixed methods study was that it was time consuming and difficult. There is an issue about what one is attempting to do. Researchers can use the whole framework over the life—and afterlife—of a study to ensure they meet the best quality standards. A user of research is more likely to want to know whether the quality of a study is “good enough.” In the field of evidence synthesis, devising a minimum rather than comprehensive set of criteria was the goal. There is a need to identify the most important criteria, or at least prioritize them. Prioritization is likely to depend on the paradigm of the decision maker. This makes the planned Delphi study mentioned earlier one of the most important next steps for the development of quality assessment in mixed methods research.

**How to assess the individual components.** Some authors argue that any quality assessment of mixed methods research must include separate evaluation of the quality of the individual components—where each component is assessed by criteria acceptable to its methodology—as well as a quality assessment of the whole mixed methods study (O’Cathain et al., 2008). Others find this use of three sets of standards cumbersome—indeed, an obstacle—and instead propose an integrated framework that can be applicable to each component as well as to the whole study (Tashakkori & Teddlie, 2008; Teddlie & Tashakkori, 2009). The latter approach is very attractive, and it remains to be seen how acceptable it is to both

qualitative and quantitative researchers working together on mixed methods studies.

**Competing criteria.** A study may meet one criterion and, by doing so, be less likely to meet another criterion. An attempt has been made to assess this when applying the framework to the example in Box 21.1. This issue did not arise but may do so when the framework is tested on more example studies.

**Is it really comprehensive?** The framework may not include the work of some mixed methods scholars because I failed to find their books or papers. When looking back at the framework now, I already have concerns about the lack of visibility of paradigms, and readers may see more gaps. This chapter is most definitely “toward” a comprehensive framework and has yet to arrive at its destination.

**Learning about quality.** The quality of mixed methods research needs to be a central part of teaching and training in this approach. Finding ways of consolidating the language of quality in mixed methods research will facilitate learning.

### ◆ Conclusions

Over the past few years, a number of mixed methods scholars have considered how best to assess the quality of mixed methods research. The work of those who have made the mixed methods aspects of a study central to their assessment has been brought together into a comprehensive framework. This framework consists of eight quality domains and is structured by the journey of a research study from planning through to data collection, interpretation, and use in the real world. It is put forward as a first attempt at a comprehensive framework in the hope that it will be developed further in the future.

### Research Questions and Exercises

1. Take the framework and apply it to your own mixed methods study. Test your understanding of each item in the framework and, at the same time, test the framework:
  - Are the domains comprehensive?
  - Is it possible to assess each item within a domain?
  - Does it help you to identify ways of improving your study?
  - Can you conclude that your study is a good or a poor mixed methods study?

You may wish to write a paper about this and publish your findings with the aim of contributing to the development of the framework and understanding of how to assess the quality of mixed methods research.

### ◆ References

- Bryman, A. (1988). *Quantity and quality in social research*. London: Routledge.
- Bryman, A. (2006). Paradigm peace and the implications for quality. *International Journal of Social Research Methodology*, 9, 111–126.
- Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research*, 1, 1–18.
- Bryman, A., Becker, S., & Sempik, J. (2008). Quality criteria for quantitative, qualitative and mixed methods research. *International Journal of Social Research Methodology*, 11, 261–276.
- Caracelli, V. J., & Riggan, L. J. C. (1994). Mixed-method evaluation: Developing quality criteria through concept mapping. *Evaluation Practice*, 15(2), 139–152.
- Chen, H. (1997). Applying mixed methods: A dominant methodology for the future? In J. C. Greene & V. J. Caracelli (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms* (pp. 61–72). San Francisco: Jossey-Bass.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). London: Sage.
- Bryman, A. (1988). *Quantity and quality in social research*. London: Routledge.
- Bryman, A. (2006). Paradigm peace and the implications for quality. *International Journal of Social Research Methodology*, 9, 111–126.
- Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research*, 1, 1–18.
- Bryman, A., Becker, S., & Sempik, J. (2008). Quality criteria for quantitative, qualitative and mixed methods research. *International Journal of Social Research Methodology*, 11, 261–276.
- Caracelli, V. J., & Riggan, L. J. C. (1994). Mixed-method evaluation: Developing quality criteria through concept mapping. *Evaluation Practice*, 15(2), 139–152.
- Chen, H. (1997). Applying mixed methods: A dominant methodology for the future? In J. C. Greene & V. J. Caracelli (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms* (pp. 61–72). San Francisco: Jossey-Bass.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). London: Sage.
- Creswell, J. W., Fetters, M. D., Plano Clark, V. L., & Morales, A. (2009). Mixed methods intervention trials. In S. Andrew & E. J. Halcomb (Eds.), *Mixed methods research for nursing and the health sciences* (pp. 161–180). Chichester, UK: Wiley-Blackwell.
- Creswell, J. W., & Plano Clark, V. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Datta, L. (1997). A pragmatic basis for mixed-method designs. In J. C. Greene & V. J. Caracelli (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms* (pp. 33–46). San Francisco: Jossey-Bass.
- Dellinger, A. B., & Leech, N. L. (2007). Toward a unified validation framework in mixed methods research. *Journal of Mixed Methods Research*, 1(4), 309–332.
- Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., & Sutton, A. (2005). Synthesising qualitative and quantitative evidence: A review of possible methods. *Journal of Health Services Research and Policy*, 10(1), 45–53.
- Erzberger, C., & Kelle, U. (2003). Making inferences in mixed methods: The rules of integration. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social*

- of behavioral research (pp. 457-488). Thousand Oaks, CA: Sage.
- Giddings, L. S., & Grant, B. M. (2009). From rigour to trustworthiness: Validating mixed methods. In S. Andrew & E. J. Halcomb (Eds.), *Mixed methods research for nursing and the health sciences* (pp. 119-134). Chichester, UK: Wiley-Blackwell.
- Harden, A., & Thomas, J. (2010). Mixed methods and systematic reviews: Examples and emerging issues. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed.). Thousand Oaks, CA: Sage.
- Karrak, P., Bialocerkowski, A. E., Massy-Westropp, N., Kumar, V. S., & Grimmer, K. A. (2004). A systematic review of the content of critical appraisal tools. *BMC Medical Research Methodology*, 4, 22.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.
- Maxwell, J. A., & Mittapalli, K. (2010). Realism as a stance for mixed method research. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed.). Thousand Oaks, CA: Sage.
- Mertens, D. (2003). Mixed methods and the politics of human research: The transformative-empiricist perspective. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 135-164). Thousand Oaks, CA: Sage.
- Miller, S. (2003). Impact of mixed methods and design on inference quality. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 423-455). Thousand Oaks, CA: Sage.
- Moher, D., Jadad, A. R., Nichol, G., Penman, M., Tugwell, P., & Walsh, S. (1995). Assessing the quality of randomized controlled trials: An annotated bibliography of scales and checklists. *Controlled Clinical Trials*, 16(1), 62-73.
- Murphy, E., Dingwall, R., Greatbarch, D., Parker, S., & Watson, P. (1998). *Qualitative research methods in health technology assessment: A review of the literature* (No. 2 (16)). Health Technology Assessment.
- O'Carthain, A., Murphy, E., & Nicholl, J. (2007). Why, and how, mixed methods research is undertaken in health services research: A mixed methods study. *BMC Health Services Research*, 7, 85.
- O'Carthain, A., Murphy, E., & Nicholl, J. (2008). The quality of mixed methods studies in health services research. *Journal of Health Services Research and Policy*, 13(2), 92-98.
- O'Carthain, A., Walters, S. J., Nicholl, J. P., Thomas, K. J., & Kirkham, M. (2002). Use of evidence based leaflets to promote informed choice in maternity care: Randomised controlled trial in everyday practice. *BMJ*, 324, 643-646.
- Onwuegbuzie, A., & Combs, J. (2010). Emergent data analysis techniques in mixed methods research: A synthesis. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed.). Thousand Oaks, CA: Sage.
- Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools*, 13(1), 48-63.
- Paterson, B. L., Thorne, S. E., Canam, C., & Jilings, C. (2001). *Meta-study of qualitative health research. A practical guide to meta-analysis and meta-synthesis*. London: Sage.
- Pluye, P., Gagnon, M., Griffiths, F., & Johnson-Lafleur, J. (2009). A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in mixed studies reviews. *International Journal of Nursing Studies*, 46(4), 529-546.
- Pluye, P., Grad, R. M., Dunkowski, L. G., & Stephenson, R. (2005). Impact of clinical information-retrieval technology on physicians: A literature review of quantitative, qualitative and mixed methods studies. *International Journal of Medical Informatics*, 74, 745-768.
- Pope, C., Mays, N., & Popay, J. (2007). *Synthesizing qualitative and quantitative health evidence. A guide to methods*. New York: McGraw Hill Open University Press.

- Sale, J. E. M., & Brazil, K. (2004). A strategy to identify critical appraisal criteria for primary mixed method studies. *Quality and Quantity*, 38(4), 351-365.
- Sandelowski, M., Voils, C. I., Knaff, G. (2009) On quantizing. *Journal of Mixed Methods Research*, 3(3), 208-222.
- Silverman, D. (2000). *Doing qualitative research. A practical handbook*. London: Sage.
- Stapleton, H., Kirkham, M., & Thomas, G. (2002). Qualitative study of evidence based leaflets in maternity care. *BMJ*, 324, 639-43.
- Steckler, A., Mcleroy, K. R., Goodman, R. M., Bird, S. T., & McCormick, L. (1992). Toward integrating qualitative and quantitative methods: An introduction. *Health Education Quarterly*, 19(1), 1-8.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. London: Sage.
- Tashakkori, A., & Teddlie, C. (2003). 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 & behavioral research* (pp. 671-701). Thousand Oaks, CA: Sage.
- Tashakkori, A., & Teddlie, C. (2008). Quality of inferences in mixed methods research. In M. Bergman (Ed.), *Advances in mixed methods research: Theories and applications*. London: Sage.
- Tashakkori, A., & Teddlie, C. (2009) Integrating qualitative and quantitative approaches to research. In L. Bicklen & D. Rog (Eds.), *The handbook of applied social research methods* (2nd ed., pp. 283-317). Thousand Oaks, CA: Sage.
- Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioural sciences. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioural research* (pp. 3-50). Thousand Oaks, CA: Sage.
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks: Sage.
- Turner-Stokes, L., Harding, R., Sergeant, J., Lupton, C., & McPherson, K. (2006). Generating the evidence base for the National Service Framework (NSF) for long term conditions: A new research typology. *Clinical Medicine*, 6, 91-97.

- Committee on Crossing the Quality Chasm: Adaptation to Mental Health and Addictive Disorders. (2006). *Improving the quality of care for mental and substance-use conditions: Quality chasm series*. Washington, DC: National Academies Press.
- Committee on Facilitating Interdisciplinary Research. (2004). *Facilitating interdisciplinary research*. Washington, DC: National Academies Press.
- Dahlberg, B., Barg, F. K., Gallo, J. J., & Wittink, M. N. (2009). Bridging psychiatric and anthropological approaches: The case of "nerves" in the United States. *Ethos*, 37, 283-313.
- Gallo, J. J., Bogner, H. R., Morales, K. H., & Ford, D. E. (2005). Patient ethnicity and the identification and active management of depression in late life. *Archives of Internal Medicine*, 165(17), 1962-1968.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine Publishing.
- Glass, T. A., & McAtee, M. (2006). Behavioral science at the crossroads in public health: Extending horizons, envisioning the future. *Social Science & Medicine*, 62, 1650-1671.
- Institute of Medicine. (2001). *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academies Press.
- 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.
- Morse, J. M. (2008). "What's your favorite color?" Reporting irrelevant demographics in qualitative research. *Qualitative Health Research*, 18(3), 299-300.
- National Institute of Mental Health. (2008). *National Institute of Mental Health strategic plan*. Washington, DC: Author.
- Stange, K. C., Crabtree, B. F., & Miller, W. L. (2006). Publishing multimethod research. *Annals of Family Medicine*, 4, 292-294.
- Switzer, J., Wittink, M., Karsch, B. B., & Barg, F. K. (2006). "Pull yourself up by your bootstraps": A response to depression in older adults. *Qualitative Health Research*, 16(9), 1207-1216.
- Tashakkori, A., & Teddlie, C. (2008). Quality of inferences in mixed methods research: Calling for an integrative framework. In M. M. Bergman (Ed.), *Advances in mixed methods research* (pp. 100-119). London: Sage.
- Tichy, H. J. (1988). *Effective writing for engineers, managers, scientists* (2nd ed.). New York: John Wiley & Sons.
- U.S. Department of Health and Human Services. (1999). *Mental health: A report of the surgeon general—older adults and mental health*. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health.
- U.S. Department of Health and Human Services. (2001). *Mental health: Culture, race and ethnicity. A supplement to mental health: A report of the surgeon general*. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.
- Wittink, M. N., Barg, F. K., & Gallo, J. J. (2006). Unwritten rules of talking to doctors about depression: Integrating qualitative and quantitative methods. *Annals of Family Medicine*, 4(4), 302-309.
- Wittink, M. N., Dahlberg, B., Biruk, C., & Barg, F. K. (2008). How older adults combine medical and experiential notions of depression. *Qualitative Health Research*, 18(9), 1174-1183.
- Wittink, M. N., Joo, J. H., Lewis, L. M., & Barg, F. K. (2009). Losing faith and using faith: Older African Americans discuss spirituality, religious activities, and depression. *Journal of General Internal Medicine*, 24(3), 402-407.
- Wittink, M., Morales, K. H., Meoni, L. A., Ford, D. E., Wang, N. Y., Klag, M. J., et al. (2008). Stability of preferences for end-of-life treatment after 3 years of follow-up: The Johns Hopkins Precursors Study. *Archives of Internal Medicine*, 168, 2125-2130.

## 31

## EPILOGUE

## Current Developments and Emerging Trends in Integrated Research Methodology

◆ Abbas Tashakkori and Charles Teddlie

*Here we are, at the end of our journey through mixed methods, or is it the beginning?*

(Tashakkori & Teddlie, 2003c, p. 671)

In the preface to this volume, we shared with you some of our experiences in what we called a "journey in time" while we worked on the two editions of the *Handbook* (Tashakkori & Teddlie, 2003a, and the current volume). It has been a challenging and enriching journey, indeed. Once again, we find ourselves simultaneously at the end and in the beginning of a fantastic and enriching journey. The second edition of the *Handbook* has been

another window to the continuously evolving landscape of integrated methodology. We would like to share with our readers some of our experiences and observations within that landscape. The mixed methods community, as it has been called by various scholars (Denscombe, 2008; Morgan, 2007; Tashakkori, 2009), has gone through a relatively rapid growth spurt. Scholars writing within the two volumes of the *Handbook*, and outside of it, have