

# RESEARCH IN THE SCHOOLS

A nationally refereed journal sponsored by the  
Mid-South Educational Research Association  
and the University of North Florida

---

Volume 13, Number 1

Spring 2006

**SPECIAL ISSUE: NEW DIRECTIONS IN MIXED METHODS RESEARCH**  
**Guest Editor, R. Burke Johnson**

Foreword . . . . .	<i>i</i>
<i>Michael Q. Patton</i>	
Foreword . . . . .	<i>iii</i>
<i>James E. McLean</i>	
Editorial: Note From the Guest Editor . . . . .	<i>v</i>
<i>R. Burke Johnson</i>	
<b><i>Philosophical Assumptions and Stances</i></b>	
How Interpretive Qualitative Research Extends Mixed Methods Research . . . . .	<i>1</i>
<i>John Creswell, Ron Shope, Vicki L. Plano Clark, and Denise O. Green</i>	
<b><i>Inquiry Logics</i></b>	
A General Typology of Research Designs Featuring Mixed Methods . . . . .	<i>12</i>
<i>Charles Teddlie and Abbas Tashakkori</i>	
Defining and Designing Mixed Methods Research Synthesis Studies . . . . .	<i>29</i>
<i>Margarete Sandelowski, Corrine I. Voils, and Julie Barroso</i>	
Mixed Methods Research: Are the Methods Genuinely Integrated or Merely Parallel? . . . . .	<i>41</i>
<i>Robert K. Yin</i>	
The Validity Issue in Mixed Research . . . . .	<i>48</i>
<i>Anthony J. Onwuegbuzie and R. Burke Johnson</i>	
<b><i>Guidelines for Practice</i></b>	
The Contribution of Computer Software to Integrating Qualitative and Quantitative Data Analyses . . .	<i>64</i>
<i>Pat Bazeley</i>	
A Theory-Driven Evaluation Perspective on Mixed Methods Research . . . . .	<i>75</i>
<i>Huey T. Chen</i>	
<b><i>Sociopolitical Commitments</i></b>	
Enhancing the Policy Process Through the Use of Ethnography and Other Study Frameworks: A Mixed-Method Strategy . . . . .	<i>84</i>
<i>Valerie J. Caracelli</i>	
<b><i>Synthesis and Future Research</i></b>	
Toward a Methodology of Mixed Methods Social Inquiry . . . . .	<i>93</i>
<i>Jennifer C. Greene</i>	

## Foreword: Trends and Issues as Context

Michael Quinn Patton

*Utilization-Focused Evaluation*

What are the top 10 trends in program evaluation today? Or the top 10 issues? Or the top 10 challenges?

The framing question can take any of these several forms when I'm asked to give a keynote speech or conference presentation or guest seminar on the status and future of program evaluation. My full top ten list varies depending on a particular audience and its interests, but the top five is pretty fixed -- and mixed methods is always in the top five. In addition to publication of the first *Handbook of Mixed Methods in Social and Behavioral Research* (Tashakkori & Teddlie, 2003) and the new *Journal of Mixed Methods Research* (JMMR), wherever I go there's a crescendo buzz about mixed methods that may prove to be a tipping point. Or may not. There are important counterforces afoot in the land, but before I address those, let me acknowledge the significance of this special issue.

*Research in the Schools* has assembled a stellar group of authors to address a variety of cutting edge issues: philosophical, conceptual, methodological, sociopolitical, and practical. These experienced and knowledgeable authors address integrating qualitative and quantitative approaches during design, data collection, and analysis; synthesizing findings; challenges of validity, credibility, and legitimization; creating a strong mix and distinguishing types of mixes; mixed methods issues in federal accountability and theory-driven evaluation; and mixed methods as a form of social inquiry. It's an impressive collection and my crystal ball says this special issue will quickly become a classic, much cited and widely disseminated.

It comes at a good time. I hinted above that there are forces running counter to the wisdom and value of mixed methods. Those forces are represented by the ongoing insistence in some quarters that randomized controlled trials are the "Gold Standard" for research and evaluation and that any other designs, including mixed methods designs, fall short. This is the position represented by the guidelines and review procedures of the Institute of Education Sciences (U.S. Department of Education), the influential *What Works Clearinghouse*, and the Campbell Collaborative, to name but a few powerful standard-setters in educational research and evaluation. These experimental design advocates describe their position as supporting *scientifically based methods*, language that is both politically

powerful and seductive. Who wants to be non-scientific? Yet their actual operational criteria give higher ratings to pure experimental designs than to mixed methods designs.

When the Institute of Education Sciences was formulating its narrow guidelines, the American Evaluation Association (AEA) adopted an official organizational policy response that said, in part:

Many methods are capable of demonstrating scientific rigor. For at least a decade, evaluators publicly debated whether newer inquiry methods were sufficiently rigorous. This issue was settled long ago. Actual practice and many published examples demonstrate that alternative and mixed methods are rigorous and scientific. To discourage a *repertoire of methods* would force evaluators backward. (AEA, n.d.)

This AEA position was ignored by the Institute of Education Sciences and proved controversial within the larger AEA membership, leading some distinguished members to drop out of AEA as a result. The debate continues both within the evaluation profession and in the larger society. At this point, it seems to me, the money and power in educational research favor the advocates of randomized controlled trials (RCTs) as the gold standard, a position that, in my judgment, is more ideologically than methodologically based. For, as this special issue makes clear, mixed methods bring particular strengths and insights to educational inquiry.

My own view, consistent with the AEA policy statement, is that the gold standard should be *methodological appropriateness* rather than methodological orthodoxy. Methodological appropriateness means that designs should be judged on the extent to which they answer the inquiry question at hand, not whether they adhere to some preordained standard. Making RCTs the gold standard puts the method before the question, a fundamental violation of inquiry in any field.

In the context of the debate about what constitutes the gold standard for educational research and evaluation, let me just note one of the strengths of mixed methods, namely the flexibility and adaptability

of such methods. To emphasize these traits, let me remind readers of the origin of the gold standard metaphor, for therein lies a cautionary tale. The gold standard is a monetary system in which the standard economic unit of account is a fixed weight of gold. When several nations are using such a fixed unit of account, the rates of exchange among national currencies effectively become fixed. The United States stopped issuing promises to redeem dollars for gold in 1933 – part of a policy change for dealing with the Great Depression. The 1944 Bretton Woods system created an obligation for each country to maintain the exchange rate of its currency in terms of gold. The system collapsed in 1971, following the United States' suspension of convertibility from dollars to gold. The gold standard failed the international economic system precisely because of its rigidity. RCTs as the gold standard in research and evaluation suffer this same rigidity. Mixed methods, within a framework that values methodological appropriateness and triangulation, offer, in contrast, flexibility and adaptability.

Thus, as you read the exquisite articles in this special issue, you will be learning about important methodological advances and pace-setting applications. You will also be participating in the larger debate that will determine the future criteria for judging research quality. By what standards should research and evaluation be judged? This special issue should go a long way in helping you decide and articulate your own views on this important matter.

Before closing, perhaps I should add one final piece. I opened by observing that, in my judgment, mixed methods is one of the top five trends in evaluation. Some readers may be wondering what the other four are. Let me relieve the suspense so that those of you reading in bed can sleep in peace. They are: (1) globalization of evaluation; (2) professionalization of evaluation; (3) logic models and theories of change as the frameworks for evaluation; and (4) systems thinking and complexity perspectives (which challenge linear logic models).

#### References

Tashakkori, A., & Teddlie, C. (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.

## Foreword: From Quantitative to Qualitative and Half Way Back

**James E. McLean**  
*University of Alabama*

As a classically trained statistician (M.Stat., 1971), I found limitations imposed by having only quantitative methodologies in my toolbox a hard pill to swallow. That realization came quickly in the early 1970s as I became more involved in evaluation. At that time, good qualitative references were scarce and most qualitative people seldom spoke with quantitative people and vice versa. In addition, most of the material on qualitative methods was from disciplines that were foreign to me at the time. However, it was obvious that different types of evaluation questions (or research questions) required different methodologies and many evaluation questions could only be addressed with qualitative methodologies. It is interesting that among my first references for qualitative methods were those of Michael Q. Patton. Indeed, my first formal training in qualitative methods was a workshop given by Michael Patton at a professional meeting in the early 1970s. Over the years, it has been refreshing to see many of the barriers come down and a new generation of researchers marrying qualitative and quantitative methods into the mixed methods approach.

While I never considered myself to be a “mixed methods” person (a term that did not exist in the 1970s), I must admit that my applications began to use a greater combination of quantitative and qualitative methods after my epiphany in the early 1970s. Fortunately, my involvement with the Evaluation Network and the Evaluation Research Society from their beginnings (1975 and 1976 respectively) allowed me the opportunity to interact regularly with many of the qualitative pioneers (e.g., Michael Patton, Egon Guba, & Yvonna Lincoln). These two groups merged to become the American Evaluation Association in 1988. By the end of the 1980s, the concept of mixed methodology was taking shape. In 1997, even the National Science Foundation was on board having published a book on mixed methods titled *User-Friendly Handbook for Mixed Method Evaluations* (Frechtling & Sharp, 1997). Beginning in January 2007, a new journal titled the *Journal of Mixed Methods Research* is scheduled for publication (Sage Publishing). This is to say that the *Research in the Schools (RITS)* special issue on mixed methods is very timely, indeed. This issue should be even more successful than the two previous special issues of *RITS* (Statistical Significance Testing, Fall 1998 and Creativity, Fall 2002), both of which were referenced

often (Please note that while I was not the guest editor for either of these special issues, I was a Co-Editor of *Research in the Schools* at the time they were published and was the author of two of the articles.). That being said, let me turn my attention to some specifics of this issue.

A special issue of a journal on a topic should cover the topic in depth as well as provide a diversity of viewpoints. First, I will address the issue of coverage. The special issue provides a philosophical rationale for the existence of mixed methods research beyond that of qualitative methods (Creswell, Shope, Plano Clark, & Green). It also provides a discussion of a number of issues surrounding the logic of mixed methods as a mode of inquiry (Teddlie & Tashakkori; Sandelowski, Voils, & Barroso; Yin; Onwuegbuzie & Johnson). The special issue provides some guidelines for the practice of mixed methodology including a description of computer software that might address mixed methods needs (Bazeley) and a theoretical rationale for combining qualitative and quantitative methods in a study (Chen). In addition, the policy implications (Caracelli) and the future of mixed methods are addressed (Greene). While all of the authors demonstrate support for the idea of mixed methods, they differ markedly in their theoretical conceptualizations, approaches, and applications. Thus, I would conclude that the special issue addresses breadth, depth, and a diversity of viewpoints.

The timing of this special issue is excellent. As noted in my earlier comments, the movement towards mixed methods has been gaining steam since the early 1970s. Like movements of the past (e.g., action research), its future is not insured. In fact, a number of recent events threaten that future. The No Child Left Behind Act of 2001 specified “scientific research” a number of times in the law. This has been interpreted to mean that the use of quantitative methods with random assignment is necessary (the so-called “gold standard”). This interpretation has been further supported by its incorporation into guidelines developed by the Institute of Educational Sciences (IES). Thus, it is very important to keep the dialog going.

A strong argument in favor of using a mixed method approach is that it could go beyond the guidelines established by the IES. Such an approach does not negate the use of “random trials” as specified

by the IES, but may not only address the efficacy of a procedure (Did it work?) using a random trials quantitative approach, the qualitative component could be used to determine why it did or did not work. Both types of information are very important depending on the outcomes of the statistical approach. Answering the “why” or “why not” question can save money by providing the information to improve an effective treatment or by identifying the problems to fix a treatment that is not effective. This argument does not even consider the contributions that the qualitative component of a mixed methods approach can bring to theory development. This special issue provides a wonderful framework for promoting a continued dialog about mixed methods and for helping researchers take the best advantage of what both qualitative and quantitative methods have to offer. It also helps us move mixed methodology one step closer to becoming a “discipline” in itself.

#### References

- Frechtling, J., & Sharp, L. (1997). *User-friendly handbook form mixed method evaluations*. Arlington, VA: National Science Foundation.

## Editorial: Note From the Guest Editor

**R. Burke Johnson**

*University of South Alabama*

Working on this special edition has been a rewarding experience. My colleague, Tony Onwuegbuzie (who has worked with me closely on this special issue), and I put together our “dream table of contents” last year, and I am proud to report that our final table of contents is virtually identical to the original table of contents. We have been privileged to work with an outstanding team of authors who are leaders in mixed research. The authors have painstakingly covered an important set of topics that provides a substantial sample of the latest thinking about mixed research.

The three major approaches to research can be broadly conceptualized by drawing a continuum and placing the term qualitative research on the left pole, the term quantitative research on the right pole, and placing the term mixed research in the middle. Although Tony Onwuegbuzie and I recommend that mixed research be based on a broad and inclusive definition (Johnson & Onwuegbuzie, 2004; Onwuegbuzie & Johnson, 2006 in this issue) with much of the area in the continuum falling into the domain of mixed research, many other authors have cogently argued for more specific definitions. I believe there is merit to both of these definitional approaches. Regardless of where methodologists place mixed research on the research continuum, it is important to recognize that the purpose of mixed research is *not* to replace qualitative research or quantitative research. There will be many times and places where qualitative research will be preferred over mixed and quantitative research, and there will be many times and places where quantitative research will be preferred over qualitative and mixed research. Qualitative research, quantitative research, and mixed research *all* have much to offer the research community. This special issue is focused on the “middle” approach (i.e., mixed research).

In the first of the nine articles in this special issue, John Creswell (co-founder and Co-Executive Editor, with Abbas Tashakkori, of the *Journal of Mixed Methods Research*) and his colleagues (Ron Shope, Vicki L. Plano Clark, and Denise O. Green) convincingly explain why mixed research often is viewed positively by qualitative researchers. They hope to increase the dialogue with qualitative researchers about the legitimacy and usefulness of mixed research. This is a dialogue that I trust will continue to grow

because, in my view, the local and experienced worlds described so vividly and passionately by qualitative researchers are equally important to social science as is the world of probabilistic generalizations (and related attempts to produce broadly applicable solutions to social problems) that is described by quantitative researchers.

In Article 2, Charles Teddlie and Abbas Tashakkori (co-editors of the landmark *Handbook of Mixed Methods in Social and Behavioral Research*) provide the most inclusive and, perhaps, the best typology of mixed research offered to date. Although it is very helpful to have more specialized and competing typologies, all of us will benefit if we become familiar with the new Teddlie/Tashakkori typology of mixed research designs. In Article 3, Margarete Sandelowski (an eminent and internationally recognized qualitative research methodologist) and her colleagues (Corrine I. Voils, and Julie Barroso) have offered an insightful and cutting-edge explanation of how to conduct mixed research syntheses. This should prove useful for dissertation students constructing their literature reviews, practicing researchers as they try to understand and integrate the past literature, and researchers who want to contribute mixed research synthesis studies to the literature (along the lines of meta-synthesis in qualitative research and meta-analysis in quantitative research). In Article 4, Robert K. Yin (whose book *Case Study Research* earned the praise of the late Donald T. Campbell) provides an important article in which he calls for integrated research questions (i.e., addressing each research question with multiple rather than single methods), and he explains his useful mixed research concepts of integrated units of analysis, integration through sample nesting, using analogous items across instruments, and analytic integration. In Article 5, Tony Onwuegbuzie and I discuss nine types of legitimation for mixed research; *legitimation* is the term we suggest be used in mixed research to refer to research quality (in qualitative research the generally agreed upon term is “trustworthiness” and in quantitative research the term is “validity”).

Articles 6 and 7 provide new guidelines for practice. In Article 6, Pat Bazeley (an internationally renowned scholar of qualitative data analysis and Associate Editor of the *Journal of Mixed Methods Research*) provides an accessible and highly useful

discussion of integrating qualitative and quantitative data through analysis. By reading this article, readers will gain specific strategies for mixed data analysis and they will increase in mixed data analysis self-efficacy because of the clarity of this article. In Article 7, Huey T. Chen (originator, along with Peter Rossi, of *Theory-Driven Evaluation*) shows how mixed research is used in conjunction with Theory-Driven Evaluation through, for example, the generation and testing of program theory (which is composed of an action model and a change model). Chen also introduces several mixed research strategies for our consideration.

In Article 8, Valerie Caracelli (one of the top evaluators with the U.S. federal government and longtime writer on mixed research) discusses the current federal push for randomized clinical/controlled trials for the purpose of accountability, and she explains how this approach can be fruitfully complemented by ethnographic research strategies. She provides many examples of federal evaluations that, in the past, have used ethnographic approaches separately and in conjunction with other methodologies.

In Article 9, Jennifer Greene (coeditor, with Valerie Caracelli, of the outstanding 1997 book *Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms*) not only reviews the other eight articles, she also introduces four domains that constitute social science research methodology. The four domains include philosophical assumptions and stances (i.e., what qualitative research methodologists Denzin and Lincoln appear to mean by the term “paradigm”; a synonym may be *philosophical paradigm* or *epistemological paradigm*), inquiry logic (i.e., what traditionally is called “methodology” and is that which “structures the inquirer’s gaze”), guidelines for practice (i.e., specific tools and procedures, the “how to” part of methodology), and sociopolitical commitments (i.e., commitments arising in relation to the specific location in society in which inquiry is situated). This article is especially helpful in showing that methodology can be viewed as a multifaceted concept. We liked Greene’s four domains so much that we decided to use them to classify the rest of the articles, as can be seen in the table of contents.

When we were working on our “dream table of contents,” Tony Onwuegbuzie and I discussed including articles on the philosophy supporting mixed research, but we decided that articles on philosophical/methodological pragmatism already were available. Therefore, if one were to use this special issue as an introduction to recent thinking about mixed research (e.g., in a course on mixed research or as a supplement to a traditional qualitative or quantitative research course), I would recommend adding readings on underlying philosophy. Two papers, focusing on the

most commonly stated philosophy supporting mixed research (i.e., pragmatism), are provided by Howe (1988) and by Johnson and Onwuegbuzie (2004). A competing paper that focuses on a relatively new and inclusive form of realism that was developed for social research is provided by Maxwell (2004).

Before finishing, I want to thank Tony Onwuegbuzie for working closely with me from the beginning to the end on visualizing and developing this special issue. I want to thank Drs. Claribel Torres, Gail Hughes, and Larry Daniel for copyediting and producing this special issue. Finally, I want to thank the editors of *Research in the Schools*, Drs. Larry Daniel and Tony Onwuegbuzie, for devoting a special issue to the important topic of mixed research and for inviting me to be the guest editor.

Most of all, our thanks go to the authors of this special issue. It is their articles that will move the field forward by their conceptualizations of mixed research, their explications of the logic of mixed research, and by their explanations of many practical strategies and techniques that will be of use to all researchers. Thanks to these authors, we have a collection of nine excellent articles on mixed methods research. Enjoy!

#### References

- Howe, K. R. (1988). Against the quantitative-qualitative incompatibility thesis, or, Dogmas die hard. *Educational Researcher*, 17(8), 10-16.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26. Retrieved July 16, 2006, from <http://aera.net/publications/?id=338>
- Maxwell, J. (2004, April). *Realism as a stance for mixed methods research*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

## How Interpretive Qualitative Research Extends Mixed Methods Research

**John W. Creswell, Ron Shope, Vicki L. Plano Clark, and Denise O. Green**  
*University of Nebraska-Lincoln*

*Recently several authors have criticized mixed methods research because it relegates qualitative research to secondary or auxiliary status, it expresses this status through experimental trials that privilege quantitative research, and it fails to employ critical, interpretive approaches to qualitative research. This paper is a response to this position, and we draw on leading qualitative and feminist researchers who advance the importance of mixed methods research. We also cite empirical mixed methods articles that give priority to qualitative research, as well as mixed methods studies that use critical interpretive approaches. Our overall argument is that qualitative research can enhance mixed methods research, and we give specific examples as illustrations.*

Mixed methods research is both a methodology and a method, and it involves collecting, analyzing, and mixing qualitative and quantitative approaches in a single study or a series of studies (Creswell & Plano Clark, in press). Recent critics of this approach to inquiry argue that it largely serves the quantitative community, it relegates qualitative research to secondary status, and it strays too far from the interpretive foundation of qualitative research (Denzin & Lincoln, 2005; Howe, 2004). This thinking might come from the association these writers sometimes appear to make between mixed methods research and the experimental orientation to educational research as discussed in the No Child Left Behind Act (2001) and in the National Research Council (2002) report on the credibility of scientific research in education. This view is a limited, inaccurate, and stereotypic perspective about mixed methods research. Fueling their concerns are also a recent emphasis on “evidenced-based” research in education, and a perceived lack of quantitative training abroad, such as in the UK (Deem, 2002).

This paper is a much-needed response to recent critics of mixed methods research, and a challenge to their stance by suggesting that qualitative research can be prominent in mixed methods research rather than compromised by it. Seen in this way, mixed methods research is compatible with qualitative

research, and through mixed methods inquiry, we have a much-needed democratizing project valuable to inquirers in the social, behavioral, and human science communities. More specifically, we will address three concerns raised by these authors: that mixed methods pushes qualitative research to secondary or auxiliary status, that this secondary status is expressed as an adjunct to a more privileged experimental trial, and that mixed methods research does not employ critical, interpretive approaches to qualitative research. To argue our case, we will draw on several qualitative researchers who advocate for mixed methods research and for the combined use of qualitative and quantitative research. We will cite works by the qualitative researcher, Jennifer Mason (2006), the nursing researcher, Margarete Sandelowski (1996), and the feminist writer, Ann Oakley (1998; 2000). We will also incorporate our own writings (Creswell & Plano Clark, in press; Creswell, Plano Clark, Gutmann, & Hanson, 2003) and provide a review of empirical mixed methods studies that favor qualitative over quantitative research. We begin by reviewing the recent criticism leveled at mixed methods research.

### *The Resistance of Qualitative Research to Mixed Methods Research*

Kenneth Howe is a philosopher of education in the social foundations of education area at the University of Colorado – Boulder. In a recent issue of *Qualitative Inquiry*, he wrote about the “auxiliary role” (Howe, 2004, p. 52) of qualitative methods in mixed methods research. He views mixed methods as helping strengthen quantitative causal relationships, and the elevation of quantitative-experimental methods to the “top of the methodological hierarchy”

---

Correspondence should be addressed to John W. Creswell, Office of Qualitative and Mixed Methods Research, 114 Teachers College Hall, Dept. of Educational Psychology, University of Nebraska-Lincoln E-mail: jrcreswell1@unl.edu.

---



(p. 53). He attributes this to several factors, such as the imposition of external standards; the lobbying by groups such as the Fordham Foundation, the Manhattan Institute, and the Heritage Foundation; the endorsement of randomized trials as the “gold standard;” the backlash to the ideological posturing of researchers who provide a social critique of practice; the signing on to “what works” by methodological fundamentalists; and to a reaction against the perceived excesses of postmodernism. Most importantly, he attributes it to his view of favoritism toward experimental research as found in the No Child Left Behind Act of 2001 and in the National Research Council (2002) report, *Scientific Research in Education* (SRE). In SRE, he finds a report that views educational research as distinct from humanistic scholarship, advances the “piling up more and more truths” as cumulative knowledge, and endorses research questions aimed at understanding causal mechanisms. Indeed, the three types of research questions, according to the SRE report, that should be asked -What is happening? Is there a systematic effect? And why or how is it happening? – involves collecting both quantitative and qualitative data, a form of “mixed-methods experimentalism” (p. 49). For Howe, it raises questions about the role of qualitative methods by inferring causal relationships. What is particularly missing in the SRE Report, Howe feels, are the assumptions of qualitative-interpretive methods that involve the inclusion and dialogue with stakeholders, the exposure of “hidden” features of education, and the value-laden nature of research.

We would not be too alarmed if Howe’s work stood in isolation. But, his critique has been endorsed and extensively cited in a qualitative book playing on the international stage of qualitative research. In Norm Denzin and Yvonna Lincoln’s *Sage Handbook of Qualitative Research* (2005), they discuss mixed methods experimentation under a section called “resistances to qualitative studies” (p.8). They, like Howe, emphasize how mixed methods research views qualitative research as filling an auxiliary role and it takes qualitative research out of its “natural home” -within the critical, interpretive framework. This natural home involves including stakeholders in the dialogue of research, makes them active participants in inquiry, and helps their silenced voices to be heard (Denzin & Lincoln, 2005).

In both Howe (2004) and Denzin and Lincoln (2005), we have a limited view of mixed methods research that underemphasizes the importance of qualitative research and casts negative light on mixed methods. They unfortunately make the assumption that qualitative research in mixed methods inquiry is always given secondary or auxiliary status. In

contrast, writers that we will cite give it a primary role, calling it “qualitatively-driven mixed methods research” (Mason, 2006, p. 9). The critics further assume that qualitative research, within a mixed methods context, reinforces this secondary status, especially in experimental research. Although some experimental writers certainly include qualitative data as adjunct, other writers see it as a major arm of the intervention trial (e.g., Sandelowski, 1996), and advance a much broader role for it. Moreover, mixed methods studies involving experimental trials is only one type of design, and ample evidence suggests a priority given to qualitative in several types of mixed methods designs (Creswell, et al., 2003). Finally, the critics assume that qualitative interpretive approaches are not found or utilized in mixed methods research. To counter this thinking, we summarize the stances of several writers who have discussed the value of interpretive frameworks in mixed methods (e.g., Oakley, 2000), and we cite a growing list of empirical studies that emphasize the interpretive frameworks.

#### *Qualitatively Driven Mixed Methods Research*

Qualitative research has much to contribute to mixed methods research, and it is this message that the British sociologist, Jennifer Mason (2006), adds to the discussion. Mason, probably best known in the international qualitative research community for her book, *Qualitative Researching* (2002), feels that mixed methods explanations can be driven by qualitative research, and, indeed, qualitative research has much to add to mixed methods explanations. Using a “qualitative logic” (p. 13), she feels that social life is not defined by either quantitative or qualitative, or by simply the macro- or the micro-approaches. Mixing methods can enhance and extend the logic of qualitative explanations about the social world. Specifically, qualitative research can help develop quantitative measures, especially when there are no measures available or change is involved, because qualitative research is holistic (considers the particulars of each case) (Mason, 2006). Qualitative research also makes context explicit in explanations, rather than “attempting to control for them or edit them out” (p. 17). Our social research should also seek “dialogic explanations” – multiple relevancies and questions held together in creative tension – goals similar to the qualitative constructivist epistemology.

#### *A Broader Role for Qualitative Data in Experimental, Intervention Studies*

In the health sciences, discussions have been underway for several years about the value of incorporating qualitative research into intervention,

experimental trials. Recently this has been the case in the most prestigious medical journals, such as *The Lancet* (Malterud, 2001), the *BMJ (British Medical Journal)* (Donovan et al., 2002), and by the guidelines established by the National Institutes of Health (1999). This trend has not been limited to medical/health research alone, however, in school psychology, a Task Force on Evidence-based Interventions, was formed in 1999 and has offered recommendations for qualitative research to strengthen and evaluate the outcomes of interventions (Nastasi & Schensul, 2005).

One important voice to emerge in this discussion has been the writings by the nursing researcher, Margarete Sandelowski, at the University of North Carolina (1996; 2000; 2003). At first glance, it might be convenient to view her work as confirming the fears of Howe/Denzin/Lincoln that qualitative research plays an “auxiliary” role in experiments. The spirit of her work, however, is to elevate the role of qualitative methods in experimental trials (Sandelowski, 1996). After noting some of the limitations of clinical trials (lack of practical significance, not attuned to individual variation, inappropriate instruments and measures), she discusses three options for the use of qualitative methods in intervention trials:

Qualitative methods may be used as components of case, small sample, and larger clinical trials of interventions, before a clinical trial is begun (in studies to ‘trial’ the trial) or after a clinical trial is completed. (p. 361)

In this statement, she advances the utility of qualitative research as an important first phase of the research, as a component within the trial, and as a follow-up to help explain the results of a trial. She proceeds to discuss the use of qualitative methods to explain individual variation, to verify outcomes, and to clarify discrepancies between the actual intervention and how participants experience it.

Our work on intervention studies suggests that Sandelowski’s 1996 framework is a useful tool to broaden and expand the ways qualitative methods can enhance an intervention trial. We have found several examples of studies that fit the before-, during-, and after-trial structure, and have begun to closely look at the reasons for incorporating qualitative data and the challenges that arise in using these designs (called the “embedded” or “nested” design in which qualitative research plays a supporting role within a larger experiment) (Creswell & Plano Clark, in press). We

have also noted that when published, these intervention studies typically are presented as two articles, one qualitative, and the second the trial itself – another indicator of the relative importance of qualitative research as a stand-alone publication. We have also developed a compendium of practices of incorporating qualitative data into intervention trials to encourage qualitative research in experiments, as shown in Table 1, and we have used Sandelowski’s (1996) framework to organize these practices.

#### *Other Qualitative Applications in Mixed Methods Research*

The emphasis on qualitative research in mixed methods designs is not limited to experimental studies, contrary to what Howe/Denzin/Lincoln suggest. One type of mixed methods design is an “exploratory sequential” design in which mixed methods research begins with a qualitative arm that often shapes the direction of the entire study (Creswell & Plano Clark, in press). In this type of design, the research begins with qualitative research, such as in depth case studies, exploratory interviews or focus groups, or detailed observations of a setting, and then is followed up by a quantitative component, such as the administration of an instrument or the conduct of a survey. The qualitative component in this type of design is clearly not an adjunct. It is also helpful to note that within any given mixed methods study, the priority (Morgan, 1998) or weight (Creswell & Plano Clark, in press) can shift to the qualitative component (e.g., a large ethnography followed by a smaller survey). Signs of this priority might include: the wording of the title, the explicit identification of a guiding worldview, the primary aim in a purpose statement, the use of more space for qualitative than quantitative in the article, or a more in depth analysis of the qualitative themes than the statistical results (Creswell & Plano Clark, in press). We have explored a number of reasons that mixed methods researchers choose to emphasize qualitative data. These include participant selection; instrument development; explaining the results of quantitative surveys; helping to explain the mechanisms behind quantitatively measured relationships among variables; exploring surprising or anomalous results, or results that were unexpected based on current theories; giving voice to different perspectives; and generating a theory or model that is grounded in the viewpoints of the participants that is subsequently tested or refined using quantitative methods (Bryman 2006; Creswell et al., 2003; Morgan 1998).

---

Table 1  
Compendium of Research Objectives for Adding Qualitative Research into Intervention Trials

---

Research Objectives for Collecting Qualitative Data

---

**Before** an Intervention Trial

- Develop an instrument for use in intervention trial (when a suitable instrument is not available)
- Develop good recruiting/consent practices for participants into a intervention trial
- Understand the participants, context, and environment so that an intervention would work (i.e., applying interventions to real-life situations)
- Document a need for the intervention
- Develop a comprehensive assessment of baseline status for comparison post-trial

**During** an Intervention Trial

- Validate the quantitative outcomes with qualitative voices of the participants
- Understand the impact of the intervention on participants (e.g., barriers/facilitators)
- Understand unanticipated participant experiences during the trial
- Identify key constructs that might potentially impact the outcomes of the trial, including changes in the sociocultural environment
- Identify resources that can aid in conducting the intervention
- Understand and depict process experienced by the experimental groups
- Check on the manipulation and implementation of procedures
- Identify mediating and moderating factors

**After** an Intervention Trial

- Understand how participants in the trial view the results
  - Revise the treatment based on participant feedback
  - Explain in more depth the quantitative outcomes (than the statistical results will allow) of a trial (e.g., underrepresented variations in the trial outcomes)
  - Determine the sustained effects of an intervention after a trial
  - Understand how the mechanisms worked in a theoretical model used in a clinical trial
  - Determine if the processes in conducting the trial had treatment fidelity
  - Assess the community/context for comparisons with baseline assessment to determine if there were unanticipated outcomes (good or bad)
-

Table 2 summarizes five studies in which the qualitative methods and procedures play a prominent role. The Table outlines the problem, the type of mixed methods design, qualitative and quantitative research methods including the type of data collected, and the reason for mixing methods. In the study by Brett, Heimendinger, Boender, Morin, and Marshall (2002) qualitative ethnography approaches were used to explore contextual influences to help shape and inform the design of an experiment. The topic of study was individual perceptions of factors that affect decisions about physical activity and diet. The preliminary ethnography played a prominent role because it was the initial intent of the research team to collect qualitative data to describe the family and social context of the decisions regarding the individuals' decision about physical activity and food choice. The data collected through interviews and in-home visits not only helped the researchers to understand the family and social context, but also provided a means to focus the research on the interests and values of the families in the study.

A second study used qualitative grounded theory to follow up on initial quantitative results. Wampold et al. (1995) used a grounded theory design to develop a model of a social science process that described the nature of social interactions of chemists in an academic setting. In this two component mixed methods study, the grounded theory study was the second component which provided a means for the researchers to explain the process that task-oriented people use in participation in social situations-an issue that arose during the initial quantitative study. While the qualitative data in this study has equal priority with the quantitative study, it is included in our exemplars because the qualitative data provided a context (see Mason, 2002) of social participation, and offered both confirming and contradictory evidence about task-oriented people that was not available in the quantitative data.

Qualitative research has also been used to determine whether qualitative data will confirm the quantitative findings. This was the case in Mactavish and Schleien's (2004) study of recreation and leisure in families that have children with disabilities. The priority given to qualitative data is evident in three ways. First, prior to the study the authors conducted a small qualitative study with several families to explore their perceptions on leisure. Second, in the methods section the authors state that the study was "grounded in a naturalistic paradigm" (p.125). Finally, while both quantitative survey data and qualitative interview data were collected, qualitative validation techniques including member checking and peer reviews were used to enhance the credibility of the overall interpretation of the data.

Qualitative data have been given primary emphasis in a study to develop a survey instrument for a large sample. Kutner, Steiner, Corbett, Jahnigen, and Barton's (1999) study of terminally ill patients receiving palliative care illustrates this approach. The study's qualitative interviews with 22 terminally ill patients provided researchers with an understanding of the information needs of terminally ill patients who were receiving palliative care. The themes and statements from the interview data had priority over the quantitative survey data in the study because it provided the researchers with critical data needed to design an instrument based on the viewpoints of the terminally ill patients.

Finally, qualitative data can be used to expand and elaborate on quantitative findings. Stoldosky and Grossman's (2000) study of how competent mathematics and English teachers adapt to changing cultural diversity illustrates how qualitative research can be used to extend quantitative surveys with in-depth case studies. The priority of the qualitative data was evident in a number of ways. This study used a rigorous case study design that included interviews and observations. The thematic analysis and cross-case analysis occupied approximately 20 pages of the study compared to approximately five pages that were devoted to the analysis of the survey data. In addition, in the analysis of the survey data, references were made to the individuals in the case study.

In summary, these studies not only illustrate ways that mixed methods researchers give emphasis to qualitative data, they also illustrate the use of rigorous qualitative research within mixed methods designs. For example, Wampold et al. (1995) included an axial coding diagram, Stoldosky and Grossman (2000) included cross-case analysis, and Brett et al. (2002) used both interview and observation data in their ethnography.

#### *Use of Interpretive Frameworks in Mixed Methods Research*

Aside from the emphasis given to qualitative research in mixed methods studies, an interpretive qualitative approach is evident in writings about mixed methods and in empirical mixed methods studies. Recall that Howe, Denzin, and Lincoln were critical of mixed methods research for not incorporating qualitative research found in interpretive, critical approaches. Interpretive research involves using issues, language, and approaches to research that empower the participants, recognize their silenced voices, honor their individual differences, and position both the researcher's and the participant's views in a historical/personal/political context (Deem, 2002). Distinct interpretive

HOW INTERPRETIVE QUALITATIVE RESEARCH EXTENDS MIXED METHODS RESEARCH

Table 2  
Mixed Methods Studies that Provide Emphasis on Qualitative Research

Authors	Topic	Mixed Method Design*	Quantitative Research Design and Methods	Qualitative Research Design and Methods	Reason for Mixing Methods
Brett et al. (2002)	Diet and physical activity intervention	Exploratory	Intervention Trial Diet and physical activity intervention	Ethnography 2 waves of interviews In-home observations	Improve an intervention design
Wampold et al. (1995)	Social skills within social settings	Explanatory	Survey 2 social skills inventories	Grounded Theory Open and axial coding interviews observations	Develop a model to explain a process
Mactavish and Schleien (2004)	Recreation and leisure in families that have children with disabilities	Explanatory	Survey Family recreation questionnaire	Thematic Analysis Interviews	Validate quantitative results
Kutner et al. (1999)	Information needs of terminally ill patients	Exploratory	Survey Information needs questionnaire developed from qualitative data	Thematic Analysis Interviews	Develop an instrument
Stoldosky and Grossman (2000)	Teachers adaptation to changes in school diversity	Triangulation	Survey Survey on teacher adaptation	Case Study Interviews Observations	Provide a means to examine trends in a national study

\**Note.* A Triangulation Design is a one-phase mixed methods study in which the researcher seeks to implement quantitative and qualitative methods during the same time frame and with equal weight. An Explanatory Design is a two-phase mixed methods study in which the researcher first collects and analyses quantitative data and then builds on the results of these data in a second phase of qualitative data collection and analysis. An Exploratory Design is also a two-phase design in which the researcher first collects and analyzes qualitative data and builds on the results of these data in a second phase of quantitative data collection and analysis (Creswell & Plano Clark, in press).

communities also exist, such as racial, ethnic, gendered, disability, and gay and lesbian communities (Denzin & Lincoln, 2005; Mertens, 2003).

We have found that a number of writers from these interpretive communities have embraced mixed

methods research. For example, feminist researchers, such as Brannen (1992), Maynard and Purvis (1994), Devine and Heath (1999), and Skeggs (2001) have highlighted the importance of both qualitative and quantitative methods (Deems, 2002). A prominent feminist qualitative researcher, Oakley, also calls for

the combination of qualitative and quantitative research. Ann Oakley is a Professor of Sociology and Social Policy at the University of London Institute of Education with long-term interests in the study of the history of methodology and in gender, the family, and health. We will draw on two of her works here – an article published in *Sociology* in 1998 and her book, *Experiments in Knowing: Gender and Method in the Social Sciences* (2000). Oakley starts with the premise that methodology is itself gendered and that the quantitative/qualitative dichotomy is an ideological representation. She goes on to say that early feminist methodology texts all celebrated qualitative methods as best suited to hearing women's accounts of their experiences. These methods included participant observation, unstructured/semi-structured interviewing, life history methods and focus groups. The feminist critique contested quantitative research on several grounds, such as the subject/object dichotomy, that the knower/researcher can be neutral and value-free, and that objectivity is actually "male subjectivity." Quantitative imitation of the natural sciences with the knower as the "expert," creates an unequal power relationship, a hierarchical situation contrary to feminism's emancipatory ideals. The use of numbers in quantitative research creates artificially controlled realities, thus advancing men's desire to dominate and to exert power over people. On the other hand, qualitative methods acknowledge multiple viewpoints, the role of values, and the subjectivities of both the researcher and those researched. Thus, qualitative research has the advantage of "thinking from caring," investigating actual practices and relations, and the "iterative attention to the details of what women say, and the forms of analysis dedicated to reproducing all of this as 'faithfully' as possible" (Oakley, 1998, p. 713).

But there are challenges with qualitative research, Oakley goes on to say, such as truth-claims of using women-only samples, of interviewing that creates a hierarchical arrangement, of compromising truth-seeking when royalties are shared with participants or participants involved in data analysis, of reactivity when researchers influence their data, and of mixed interpretations when data conflict. For feminist research and advancing an emancipatory project, Oakley sees qualitative and quantitative research on a continuum rather than as a dichotomy, a continuum reinforced by a close study of the history of the social sciences which does not convey a straightforward picture of two communities of scientists. Quantitative research has served the goals of feminism, such as the large-scale social surveys, and the studies of women vis-à-vis men in the labor market, the home, and in domestic relations.

Quantitative methods and statistics have established the gendering of structural inequalities in most societies. In sum, Oakley's position is that the construction of quantitative and qualitative methods as oppositions impedes critical thinking about creating an appropriate knowledge for women. The traditional stance of feminism toward embracing qualitative as the only approach does not further feminist social research.

Oakley's comments remind us of the discussion by Reichardt and Cook in 1979 who advanced ten different ways that quantitative and qualitative research are similar, not different. However, she takes their argument one step further by suggesting the advantages of quantitative research to feminists who are primarily oriented toward using qualitative research. Unfortunately, Oakley does not provide the specifics as to how mixed methods can inform emancipatory research, and her argument lies largely in offering a critique of both quantitative and qualitative research. In all fairness, Oakley's views may be related to her work in evidence-based practices in health and education, as well as her work in the history of science (Deem, 2002).

Others, however, have more directly related emancipatory aims and theoretical frameworks in qualitative research to mixed methods inquiry. In terms of interpretive approaches in mixed methods research, Greene and Caracelli (1997) advocated for making advocacy worldviews explicit in mixed methods studies, Creswell et al. (2003) suggested that one type of mixed methods design included a "transformational" model, employing a theoretical framework (e.g., feminism), and Mertens (2003), a disability researcher, linked many phases of research (e.g., stating the research question) to the study of underrepresented populations. We have also found a number of studies that incorporated critical qualitative interpretive approaches into their inquiries that report both qualitative and quantitative data, such as Skeggs' (1999) study of gay and lesbian sexuality and violence in urban public spaces, Ely's (1995) study of women's gender construction of identity; Wajcman's (1998) study of female and male senior managers in five high technology multi-national companies; Bhopal's (2000) study of gender, race, and power of South Asian women in London; and Watkins' (1998) study of mentoring of African Americans. These are all good examples of the use of interpretive, theoretical frameworks in a mixed methods study.

#### *Turning it Around - How Can Mixed Methods Enhance Qualitative Research?*

The example studies presented in Table 2 and discussions of emancipatory frameworks for mixed

methods research present clear arguments of how the ideals and foundations of qualitative research can play prominent roles within mixed methods research. To carry this argument further, we suggest that there are also ways in which mixed methods research might also enhance the goals of qualitative research. We believe there are circumstances where quantitative data can play a supportive role to qualitative frameworks. For example, quantitative data might be the best approach to guide purposeful sampling strategies, such as guiding maximal variation sampling or theoretical sampling based on individuals' attributes, attitudes, or behaviors. Quantitative data can also be useful to more fully describe the social and historical contexts of a case, such as providing statistics regarding the AIDS epidemic within a case study of AIDS-affected families, or providing a richer description of a case, such as including depression and anxiety scores in addition to qualitative descriptions. There are times when quantitative data may further advocacy-related goals, by including the perspectives of a greater variety of individuals or being able to highlight emancipatory concerns to larger audiences, including those who value numbers. Quantitative data can also be a useful supplement to qualitative research for the purposes of theory generation, which the quantitative data might help refine a model based on larger numbers than could be feasibly interviewed. Unfortunately, we have found few published mixed methods studies that incorporate supplemental quantitative data to enhance qualitative research, but we fully expect these numbers to grow as more qualitative researchers consider how mixed methods research can better address some of their research goals.

#### Implications and Conclusions

What are the implications of our reaction to the Howe, Denzin, and Lincoln commentary on mixed methods research? We see important implications for mixed methods writers, those conducting mixed methods research, readers of mixed methods studies, and educational researchers. For those writing about mixed methods studies, we need to continue to educate writers about the literature of mixed methods (the designs, the potential experimental uses, and the employment of interpretive, theoretical frameworks). We also need to encourage researchers to include interpretive frameworks in their mixed methods studies. Mertens' (2003) chapter is a start toward developing a better understanding of incorporating emancipatory aims into all aspects of a mixed methods study.

For those conducting mixed methods research, consider the important role of qualitative research in mixed methods research. It might weigh heavily into the study as a major priority. It might begin a study, provide the context necessary in a study, explore variables and constructs that are unknown, and develop themes necessary to study underrepresented populations. It should be conducted with rigor and using the methods and procedures of qualitative research. For those reading mixed methods research, recognize that some qualitative researchers will be threatened by mixed methods research, see it primarily as a quantitative orientation toward research, and seek to keep qualitative research "pure" without being diluted by quantitative research. But, with increasing frequency, qualitative researchers are involved in the development of mixed methods research, and it has a major role in this form of research. As strong, knowledgeable qualitative researchers engage in mixed methods research, qualitative inquiry will continue to hold prominent positions in mixed methods approaches. Also, recognize that it is possible to use an interpretive, critical, theoretical framework within a mixed methods study. The examples cited in this discussion attest to it. For educational research, qualitative research can enhance mixed methods (and vice versa). We would advise, however, that the possibilities for qualitative data within experiments needs to be expanded (see Table 1) and that the three questions advanced by SRE study actually embrace a mixed methods approach to research rather than the total exclusion and minimization of qualitative research. Also, in educational research, the full array of types of mixed methods studies can be found in which qualitative data are combined with surveys, correlational, or single case designs, as well as experiments.

In this discussion, we have presented evidence to suggest that qualitative research can assume a major role in mixed methods studies and it has specific features that make it attractive to mixed methods studies. Evidence shows that qualitative research is not always in a supportive, auxiliary role to quantitative research as suggested by Howe, Denzin, and Lincoln would suggest. This is not the case in either traditional constructivist forms of qualitative research and in the more recent, interpretive, critical approaches. We have drawn on several qualitative authors who have advanced mixed methods research and see qualitative research as not only contributing to the inquiry, but also providing understanding for the research, by helping to develop intervention trials, and helping to advance an emancipatory agenda. Writers such as Oakley, Mason, and Sandelowski all

contribute their thoughts to suggest the importance of qualitative research to mixed methods and vice versa. To their voices, we add our own, through our work on a compendium of practices for incorporating qualitative data into intervention trials, and through our discussion of mixed methods designs that include those that place a priority on qualitative research as well as emphasize critical, emancipatory frameworks. Mixed methods studies that employ interpretive, critical frameworks are available and being published in the literature. Qualitative research plays an important role in explaining the social world, and it can enhance, even “drive” mixed methods research, extend experimental applications, and further emancipatory aims. This evidence should cause critics to pause and further reflect on the important role for qualitative research in mixed methods inquiry.

What explains their position may be related to a lack of training and skills in quantitative research, a perceived threat by the federal government and its reports, an attempt to keep qualitative research “pure” (Rossman & Wilson, 1985), or simply a lack of understanding of the literature and research in mixed methods. Although Howe/Denzin/Lincoln refer to methods of using qualitative data in experimental trials, their concerns may be more related to paradigms and the mixing of paradigms than the actual methods. Granted, the field of mixed methods is dispersed across the social and human sciences, and writers in the field have not done an adequate job in conveying the breadth of mixed methods studies. Added to this is that some quantitative researchers have been supportive of mixed methods research. Also, some mixed methods researchers might convey a more quantitatively-oriented than qualitatively-oriented approach to mixed methods research, such as the inclusion of explanatory data analysis (Tashakkori & Teddlie, 1998), the emphasis on standards for inferences and validity (Erzberger & Kelle, 2003), and the trend toward the slowly emerging emancipatory mixed methods studies (Creswell & Plano Clark, in press). Despite these caveats, the “ominous development” of qualitative methods as playing an auxiliary role in mixed methods experiments that Howe (2004) refers to may be that some qualitative researchers have not recognized how mixed methods can enhance the development of qualitative research. This process begins by recognizing the primary role of qualitative research in many mixed methods studies, by viewing the expanded list of design possibilities that give support to this primary role, and to the emerging use of interpretive frameworks.

The lead editors for this article were R. Burke Johnson and Anthony J. Onwuegbuzie.

#### References

- Bhopal, K. (2000). Gender, 'race' and power in the research process: South Asian women in East London. In C. Truman, D. M. Mertens, & B. Humphries (Eds.), *Research and inequality* (pp. 67-79). London: UCL Press.
- Brannen, J. (1992). *Mixing methods: Qualitative and quantitative research*. Aldershot: Avebury.
- Brett, J. A., Heimendinger, J., Boender, C., Morin, C., & Marshall, J. A. (2002). Using ethnography to improve intervention design. *American Journal of Health Promotion, 16*, 331-340.
- Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research, 6*(1), 97-113.
- Creswell, J. W., & Plano Clark, V. L. (in press). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Creswell, J. W., Plano Clark, V. L., Gutmann, M., & Hanson, W. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209-240). Thousand Oaks, CA: Sage.
- Deem, R. (2002). Talking to manger-academics: Methodological dilemmas and feminist research strategies. *Sociology, 36*, 835-855.
- 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* (pp. 1-32). Thousand Oaks, CA: Sage.
- Devine, F. & Heath, S. J. (1999). *Sociological research methods in context*. London: Macmillan.
- Donovan, J., Mills, N., Smith, M., Brindle, L., Jacoby A., Peters, T., et al. (2002). Improving design and conduct of randomized trials by embedding them in qualitative research: ProtecT (prostate testing for cancer and treatment) study. *BMJ, 325*, 766-760.
- Ely, R. J. (1995). The power in demography: Women's social constructions of gender identity at work. *The Academy of Management Journal, 38*, 589-634.
- 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 and*



- behavioral research* (pp. 457-488). Thousand Oaks, CA: Sage.
- Greene, J. C., & Caracelli, V. J. (Eds.). (1997). Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms. *New Directions for Evaluation, 74*. San Francisco: Jossey-Bass Publishers.
- Howe, K. R. (2004) A critique of experimentalism. *Qualitative Inquiry, 10*(1), 42-61.
- Kutner, J. S., Steiner, J. F., Corbett, K. K., Jahnigen, D. W., & Barton, P. L. (1999). Information needs of terminal illness. *Social Science and Medicine, 48*, 1341-1352.
- 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*, 123-141.
- Malterud, K. (2001). The art and science of clinical knowledge: Evidence beyond measures and numbers. *The Lancet, 358*, 397-400.
- Mason, J. (2002). *Qualitative researching*. London: Sage.
- Mason, J. (2006). Mixing methods in a qualitatively driven way. *Qualitative Research, 6*(1), 9-25.
- Maynard, M., & Purvis, J. (Eds.) (1994). *Researching women's lives from a feminist perspective*. London: Taylor and Francis.
- Mertens, D. M. (2003). Mixed methods and the politics of human research: The transformative-emancipatory perspective. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 135-164). Thousand Oaks, CA: Sage.
- Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative Health Research, 8*, 362-376.
- Nastasi, B. K., & Schensul, S. L. (2005). Contributions of qualitative research to the validity of intervention research. *Journal of School Psychology, 43*, 177-195.
- National Institutes of Health. Office of Behavioral and Social Sciences Research. (1999). *Qualitative methods in health research: Opportunities and considerations in application and review*. Washington D.C.: National Institutes of Health.
- National Research Council. (2002). *Scientific research in education*. Washington D. C.: National Academy Press.
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).
- Oakley, A. (1998). Gender, methodology and people's ways of knowing: Some problems with feminism and the paradigm debate in social science. *Sociology 32*, 707-732.
- Oakley, A. (2000). *Experiments in knowing: Gender and method in the social sciences*. Cambridge: Polity Press.
- Reichardt, C. S., & Cook, T. D. (1979). Beyond qualitative versus quantitative methods. In T. D. Cook & C. S. Reichardt (Eds.), *Qualitative and quantitative methods in evaluation research* (pp. 7-32). Beverly Hills, CA: Sage.
- Rossmann, G. B., & Wilson, B. L. (1985). Numbers and words: Combining quantitative and qualitative methods in a single large-scale evaluation study. *Evaluation Review, 9*, 627-643.
- Sandelowski, M. (1996). Using qualitative methods in intervention studies. *Research in Nursing & Health, 19*, 359-364.
- Sandelowski, M. (2000). Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in Nursing & Health, 23*, 246-255.
- Sandelowski, M. (2003). Tables or tableaux? The challenges of writing and reading mixed methods studies. In A. Tashakkori and C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 321-350). Thousand Oaks, CA: Sage.
- Skeegs, B. (1999). Matter out of place: Visibility and sexualities in leisure spaces. *Leisure Studies, 18*, 213-232.
- Skeegs, B. (2001). Feminist ethnography. In P. Atkinson, A. Coffey, S. Delamont, J. Lofland, & L. Lofland (Eds.), *Handbook of ethnography* (pp. 426-442). London: Sage.
- Stodolsty, S. S., & Grossman, P. L. (2000). Changing students, changing teaching. *Teacher's College Record 102*(1), 125-172.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.
- Wajcman, J. (1998). *Managing like a man*. Cambridge: Polity.
- Wampold, B. E., Ankarlo, G., Mondin, G., Marcelo, T-C., Baumler, B., & Prater, K. (1995). Social skills and social environments produced by different Holland types: A social perspective on person-environment fit

models. *Journal of Counseling Psychology*,  
42, 365-379.

Watkins, G. H. (1998, November). *Satisfaction and mentoring: An African American perspective*. Paper presented at the Annual Meeting of the Association for the Study of Higher Education, Miami, FL. (ERIC Document Reproduction Services No. ED 427596)

## A General Typology of Research Designs Featuring Mixed Methods<sup>1</sup>

**Charles Teddlie**

*Louisiana State University*

**Abbas Tashakkori**

*Florida International University*

*This article presents a general typology of research designs that features those utilizing mixed methods. The Methods-Strands Matrix includes both monomethod and mixed methods designs, but the emphasis is on the more complex and adaptable mixed ones. The article starts with a brief discussion of why typologies of mixed methods designs are valuable at this time. The Methods-Strands Matrix is produced by crossing number of methods employed (monomethod, mixed methods) by number of research strands (single, multiple). The multistrand, mixed methods cell in the matrix includes four families of MM designs: sequential, concurrent, conversion, and fully integrated. Examples of each of these strands are presented. Quasi-mixed designs are also described in which qualitative and quantitative data are collected, but there is no true integration of the findings or inferences for the overall study. We conclude that it is impossible to create a complete taxonomy of mixed methods designs, because they have an evolving nature that can spin off numerous permutations. The article concludes with a seven-step process for selecting the most appropriate mixed methods design for a research study.*

This article presents the Methods-Strands Matrix, which is a general typology of designs used in the social and behavioral sciences. This typology includes monomethod designs (qualitative or quantitative), but it features mixed methods (MM) research designs, especially four families of MM designs: sequential, concurrent, conversion, and fully integrated. Before discussing the matrix, we will briefly discuss (a) why typologies of MM research designs are useful, and (b) the dimensions that have been used by various authors to create MM design typologies.

### *Are Typologies of Mixed Methods Designs Necessary?*

Scholars writing in the field of MM research have developed typologies of mixed designs from the time the field emerged. For instance, Greene, Caracelli, and Graham (1989) examined a large number of MM studies and developed a typology for the designs used in those studies based on their design characteristics and functions.

Why have so many of their colleagues followed the lead of Greene et al. (1989) in developing MM typologies? Following are five reasons why typologies are important in MM research:

1. Typologies help researchers decide how to proceed when designing their MM studies. They provide a variety of paths, or ideal design types, that may be chosen to accomplish the goals of the study.
2. Typologies of MM research designs are useful in helping to establish a common language for the field. For instance, Morse's (1991, 2003) typology of MM research designs includes notations and abbreviations still used today.
3. Typologies of MM designs help to provide the field with an organizational structure. At this point in time, given the range of existing MM typologies, it is more accurate to say that such typologies provide the field with multiple alternative organizational structures.
4. Typologies of MM designs help to legitimize the field because they provide examples of research designs that are clearly distinct from either quantitative (QUAN) or qualitative (QUAL) research designs.
5. Typologies are useful as a pedagogical tool. A particularly effective teaching technique is to present alternative design typologies and then have the students discuss their strengths and weaknesses.

---

Correspondence for this article should be addressed to Charles Teddlie, Dept. Educational Leadership, Research, and Counseling, 113 C Peabody, Louisiana State University, Baton Rouge, LA 70803. Email [edtedd@lsu.edu](mailto:edtedd@lsu.edu)

---

*Can a Typology of MM Designs Be Exhaustive?*

While typologies of MM designs are valuable, researchers should not expect them to be exhaustive. This is an important point, especially because many researchers from the QUAN tradition expect an exhaustive “menu” of designs from which to select the “correct” one for their studies (e.g., Shadish, Cook, & Campbell, 2002).

We argue throughout this article that methodologists cannot create a complete taxonomy of MM designs, due to their (the designs’) capacity to mutate into other diverse forms. Similarly, Maxwell and Loomis (2003) concluded that “the actual diversity in mixed methods studies is far greater than any typology can adequately encompass” (p. 244). This diversity in MM designs is produced by two factors:

1. The QUAL component of MM research studies. MM research utilizes an emergent strategy in at least the QUAL component of the design. Emergent designs may evolve into other forms as QUAL data collection and analysis occur (e.g., Lincoln & Guba, 1985; Patton, 1990, 2002).
2. The opportunistic nature of MM design. In many cases, a MM research study may have a predetermined research design, but new components of the design may evolve as researchers follow up on leads that develop as data are collected and analyzed. These “opportunistic” designs may be slightly different from those contained in previously published typologies of MM designs.

Given this diversity in MM designs, our typology will feature families of designs, each of which may have several different members.

*Criteria Used in MM Research Typologies*

Table 1 presents seven criteria that authors have used to create their MM typologies (e.g., Creswell, Plano-Clark, Gutmann, & Hanson, 2003; Greene & Caracelli, 1997; Greene et al., 1989; Johnson & Onwuegbuzie, 2004; Morgan, 1998; Morse 1991, 2003). These criteria include:

- number of methodological approaches used
- number of strands or phases
- type of implementation process

- stage of integration of approaches
- priority of methodological approach
- function of the research study
- theoretical perspective

Our typology utilizes the first four criteria<sup>2</sup> in the generation of what we call the Methods-Strands Matrix. We do not use the other three criteria in our typology, which focuses on methodological components of research designs. Specific reasons for the non-inclusion of the other criteria include the following:

- The priority of methodological approach. While an important consideration, the relative importance of the QUAL or QUAN components of a research study cannot be completely determined before the study occurs. In the real world, a QUAN + qual study may become a QUAL + quan study if the qualitative data become more important in understanding the phenomenon under study, and vice versa. Because the actual priority of approach (QUAL, QUAN) is determined after the study is conducted, it is not part of our design typology.
- The function of the research study. In our opinion, the intended function of a research study (e.g., triangulation, complementarity) is not a design issue, but is related to the function that the results from the study eventually serve (e.g., to corroborate findings, to enhance or elaborate findings). Because the outcomes of a MM study come after its design, we do not include this criterion in our design typology.
- The theoretical perspective. Some analysts include theoretical perspective, such as the transformative-emancipatory orientation, as a design component. While this is an important axiological (or values) component for doing research, it is a purpose (i.e., to create social justice) of the research study, rather than a design component. For researchers working within the transformative-emancipatory orientation, the pursuit of social justice is not a design choice; rather, it is the reason for doing the research, which supersedes design choices. Therefore, we do not include theoretical perspective as a criterion in our design typology.

Table 1  
*Criteria Used in MM Research Typologies and the Design Questions They Answer*

Criterion Used	What design questions does this criterion answer?	What possible values for the criterion exist?	Is this criterion used in our typology?
(1) Number of Methodological Approaches Used	Will the study involve one method (QUAN or QUAL) or both (QUAL and QUAN)?	* Monomethods Study * Mixed Methods Study	Yes
(2) Number of Strands or Phases	Will the study involve one phase only or multiple phases?	* Monostrand * Multistrand	Yes
(3) Type of Implementation Process	Will the QUAN and QUAL data collection occur sequentially or concurrently? Will data conversion occur?	* Concurrent * Sequential * Conversion * Combination	Yes
(4) Stage of Integration of Approaches	Will the study be mixed (QUAL, QUAN) in the experiential stage only, or across stages, or other combinations?	* Across all stages * Within experiential stage only * Other combinations	Yes, but only to allow the inclusion of quasi-mixed designs
(5) Priority of Methodological Approach	Does the QUAL or QUAN component have priority, or are they equal in importance, at the onset of the study?	* QUAL+quan * QUAN+qual	No
(6) Functions of the Research Study	Which of the following functions does the research design serve?	* Triangulation * Complementarity * Development * Initiation * Expansion * Other functions	No
(7) Theoretical Perspective	Will the design be driven by a particular theoretical perspective (typically the transformative perspective)?	* Some Variant of the Transformative Perspective * No Theoretical Perspective in the Design	No

One of the reasons for limiting the number of dimensions in our typology is that it could become overly complex otherwise. As observed by Mertens (2005), sub-types can easily be constructed and/or modified within the general types, depending on the purpose of the research and the research questions. For example, within a sequential design, sub-types may be constructed on the basis of priority (Creswell et al., 2003) of either the QUAL and QUAN strand.

We should also mention that in our current classification, we have abandoned (or de-emphasized) components of our previous typologies (Tashakkori & Teddlie, 1998, 2003) for a specific reason. Newer conceptualizations of mixed methods research all recognize the fact that a study is not considered mixed

if there is no integration across stages. Previously, we have distinguished mixed methods (studies using two types of data and their analysis) from mixed models (studies that are mixed throughout, with two types of questions, data and interpretations). More recent definitions and conceptualizations in the field define mixed methods very similarly to our mixed models. We are assuming that all properly defined mixed methods studies these days are of this sort; therefore a distinction between mixed methods and mixed models studies is no longer necessary.<sup>3</sup>

Supporting this change in terminology is the definition of mixed methods in the *Call for Papers of the Journal of Mixed Methods Research* (Sage Publications, first issue expected in January 2007):

“mixed methods research is defined as research in which the investigator collects and analyzes data, integrates the findings, and draws inferences using both qualitative and quantitative approaches or methods in a single study or program of inquiry.”

Following this conceptualization, one might call the studies in which two types of data are collected, but no integration of the findings/inferences occurs as *quasi-mixed designs*. We continue the use of “stage of integration” as a fourth criterion in our typology (see Table 1) in order to allow for the recognition of these *quasi-mixed designs*.

The Methods–Strand Matrix

The approach to classifying MM research designs presented in this article has evolved over time (e.g., Tashakkori & Teddlie, 1998, 2003; Teddlie & Tashakkori, 2005, forthcoming). We employ Morse’s basic notational system, but we have developed our own set of MM research designs, which are included in the Methods-Strands Matrix. The latest incarnation of this typology was developed for three reasons:

- to more specifically locate MM designs within the larger framework of a general typology of research designs in the social and behavioral sciences
- to provide more clarity on how specific MM research designs are generated from a technical perspective that features methodological components
- to present an alternative perspective on MM research designs that features methodological components, rather than theoretical perspectives and research purposes or functions

*Generation of the Methods-Strands Matrix*

A simplified version of our typology is presented in Table 2, which is a matrix created by crossing two basic design dimensions:

- Type of approach or methods employed in the study (Monomethod or Mixed Methods).
- Number of strands (or phases) of the study (Monostrand or Multistrand).

Table 2

*The Methods-Strands Matrix: A Typology of Research Designs Featuring Mixed Methods*

Design Type	Monostrand Designs	Multistrand Designs
Monomethod Designs	<b>Cell One</b> Monomethod Monostrand Designs: (1) Traditional QUAN design (2) Traditional QUAL Design	<b>Cell Two</b> Monomethod Multistrand Designs: (1) Concurrent Monomethod a. QUAN+QUAN b. QUAL+QUAL (2) Sequential Monomethod a. QUAN→QUAN b. QUAL→QUAL
Mixed Methods Designs	<b>Cell Three</b> Quasi-Mixed Mono-Strand Designs: Monostrand Conversion Design	<b>Cell Four</b> A) Mixed Methods Multistrand Designs: (1) Concurrent Mixed Designs (2) Sequential Mixed Designs (3) Conversion Mixed Designs (4) Fully Integrated Designs B) Quasi-Mixed Multi-Strand Designs: Designs Mixed at the Experiential Stage Only, including the Concurrent Quasi-Mixed Design <sup>4</sup>

It is more accurate to state that the matrix contains “families” of research designs, because each of its four cells includes numerous designs.

*Four Decision Points in the Methods-Strands Matrix*

Investigators make four basic methodological decisions when selecting a design for their study from the matrix (see Table 1, Criteria 1-4). The first two decision points create the matrix itself, which crosses the number of methodological approaches used and the number of strands (or phases) of a study.

*Number of Methodological Approaches Used.*

The Methods-Strands Matrix presented in Table 2 conceptually encompasses all three research approaches (QUAL, QUAN, MM), because it also includes purely QUAN and QUAL designs. The emphasis in this article is on the MM designs, but it is also useful to consider how they are conceptually related to monomethod QUAL and QUAN designs. The basic definitions of monomethod and MM designs are as follows:

- Monomethod designs – a type of research design in which only the QUAL approach, or only the QUAN approach, is utilized across all stages of the study.
- Mixed methods designs – a type of research design in which QUAL and QUAN approaches are mixed across the stages of a study.

*Number of Strands or Phases in the Research Design.* The second dimension of the Methods Strands Matrix presented in Table 2 refers to whether the research study has only one strand or more than one strand. The basic definitions of these terms are as follows:

- Strand of a research design – is a phase of a study that includes three stages: the conceptualization stage, the experiential stage methodological/analytical, and the inferential stage.
- Monostrand design - these designs employ only a single phase and it encompasses all of the stages from conceptualization through inference.
- Multistrand design - these designs employ more than one phase<sup>5</sup>; there are multiple phases to the study and each encompasses all of the stages from conceptualization through inference.

Following are definitions regarding stages:

- Stage – refers to a step or component of a strand/phase of a study
- Conceptualization stage – the sphere of concepts (abstract operations), which includes

the formulation of research purposes, questions, etc.

- Experiential (methodological/analytical) stage – the experiential sphere (concrete observations and operations), which includes methodological operations, data generation, analysis, etc.
- Inferential stage – the sphere of inferences (abstract explanations and understandings), which includes emerging theories, explanations, inferences, etc. (Tashakkori & Teddlie, 2003, p. 681)

A simplified outline of the strand-stage terminology is presented in Figure 1 in which there is one strand (QUAL or QUAN) in a monomethod design, with three stages:

1. Conceptualization stage
2. Experiential (methodological/analytical) stage
  - a. Methodological
  - b. Analytical
3. Inferential stage

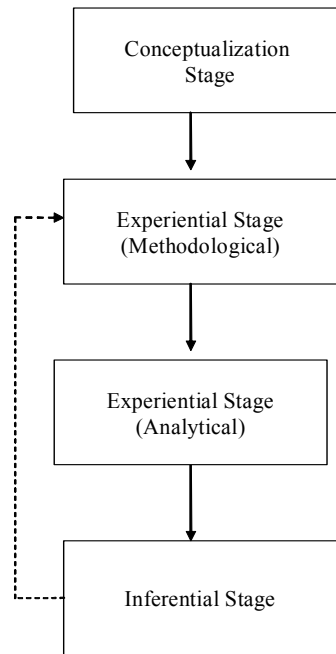


Figure 1  
Graphic Presentation of Traditional QUAN or QUAL Designs (Monomethod Monostrand design)

Dividing a strand into distinct stages allows for the transformation of one methodological approach to another within a strand; that is, a strand might start out as a QUAL, but then become QUAN, or vice versa.

*Type of Implementation Process (Concurrent, Sequential, Conversion).* This decision point involves answers to two questions: “Will the QUAN and QUAL data collection occur sequentially or concurrently? Will data conversion occur?”

Concurrent and sequential designs been employed by numerous authors writing in the MM tradition (e.g., Creswell et al., 2003; Johnson & Onwuegbuzie, 2004; Morse, 1991, 2003). In concurrent designs, the strands of a study occur in parallel or synchronous manner, whereas in sequential designs they occur in chronological order with one strand emerging from the other.

The third way to implement a MM study is to use a conversion design. These designs are a unique feature of MM research and include the following terms:

- Data conversion (transformation): collected QUAN data types are converted into narratives that can be analyzed qualitatively, and/or QUAL data types are converted into numerical codes that can be statistically analyzed.
- Quantitizing: converting QUAL data into numerical codes that can be statistically analyzed (e.g., Miles & Huberman, 1994).
- Qualitizing: the process by which QUAN data are transformed into data that can be analyzed qualitatively (e.g., Tashakkori & Teddlie, 1998).

Following the discussion above, if there is no integration of both types of analysis (i.e. if only QUAN or QUAL type of analysis is performed on the converted data), the design should be considered a *quasi-mixed design*.

*Stage of Integration of Approaches.* The final decision point is the least important for most researchers: does the integration of approaches occur in the experiential (methodological/analytical) stage only, or does it occur across stages, or other combinations? The most dynamic and innovative of the MM designs contained in the matrix are mixed across stages, so why should we consider selecting a design that is mixed in only the experiential stage?

The answer is simple—because these designs have been described in the literature and because at least one of them (what we label the Concurrent Quasi-Mixed Design in Table 2) is a popular one. In these designs, researchers who are working within one approach primarily (e.g., the QUAN approach) might elect to gather and analyze data associated with the other approach (e.g., the QUAL approach) in order to triangulate data sources, or in order to answer different aspects of the same research question. If the design is “mixed” only in the methodological/analytical stage of

the study, without deliberate integration, it should be called a *quasi-mixed design*.

#### Designs in The Methods-Strands Matrix

The remainder of this article presents some of the more important and widely used designs from the matrix, together with some illustrative diagrams. The illustrative diagrams were first presented in Tashakkori and Teddlie (2003), p. 684-690. The general features of these figures are presented in Box 1.

#### Box 1

##### The General Features of Figures 1-7

Rectangles and ovals represent either a QUAL or a QUAN stage of a research strand. If the stages are all rectangles, this indicates that the figure represents a monomethod design. If some of the stages are rectangles and some are ovals, this indicates that the figure represents a MM design.

Each strand found in the figures has three stages (conceptualization, experiential, inferential). The experiential stage is broken into two parts (methodological and analytical) to allow for conversion designs. We have divided the experiential stage into two parts methodological/analytical on all figures for the sake of consistency.

There is a broken line arrow in each figure going from the inferential stage to the methodological stage. This indicates that conclusions emerging from the inferential stage of a study may lead to the gathering of further data and further analysis in the same study. The methodological-analytical-inferential loop of each diagram is iterative.

These figures were first presented in Tashakkori and Teddlie (2003), pp. 684-690.

#### *Monomethod Designs*

There are two types of monomethod research designs, those with only one strand (Cell One in Table 2) and those with more than one strand (Cell Two in Table 2). Cell One designs are Monomethod Monostrand designs, while Cell Two designs are Monomethod Multistrand designs.

*Monomethod Monostrand Designs.* Cell One designs use a single research method or data collection technique (QUAN or QUAL) and corresponding data analysis procedures to answer research questions employing one strand only. This strand may be either QUAN or QUAL, but not both. All stages within the strand (conceptualization, experiential, inferential) are



consistently either QUAN or QUAL. (Refer to Figure 1 for an illustration of Monomethod Monostrand Designs.) Figure 1, presented in a previous section of this article, is an example of such a design.

These designs appear to be the simplest of those presented in Table 2, yet they can be quite complex (e.g., multilevel QUAN designs, detailed ethnographic QUAL designs) and have been written about in numerous books on QUAN and QUAL design. For example, the methodology for ethnographic studies has been discussed in detail in several texts (e.g., Fetterman, 1998; LeCompte & Preissle, 1993).

We will present only one example here, because our focus is on the MM designs in Cells Three and Four. Among the most well known of the QUAN designs in the social and behavioral sciences are the quasi-experimental designs first presented by Campbell and Stanley (1963) and revised in later texts (Cook & Campbell, 1979; Shadish et al., 2002). These authors used a simple notational system in which treatments were designated as X, observations were designated as O (e.g., O<sub>1</sub>, O<sub>2</sub>), R represented random assignment to treatment, and a dashed line between groups (-----) indicated nonrandom assignment to treatment. For example, the following quasi-experimental design was presented in Campbell and Stanley (1963) as the nonequivalent control group design:

O <sub>1</sub>	X	O <sub>2</sub>
O <sub>1</sub>		O <sub>2</sub>

When quasi-experimental studies involve the collection of QUAN data only, they are examples of Monomethod Monostrand Designs.

*Monomethod Multistrand Designs.* Cell Two designs employ a single method or data collection technique (QUAN or QUAL) and corresponding data analysis procedures to answer research questions. These designs use two or more strands, which may be either QUAN or QUAL, but not both. There are two types of Monomethod Multistrand Designs:

- **Concurrent Monomethod Multistrand Designs.** In Cell Two of Table 2 these designs are designated as QUAN + QUAN or QUAL + QUAL. In these designs, there are multiple strands of the research design which occur in a parallel manner. All stages within the strands (conceptualization, experiential, inferential) are consistently either QUAN or QUAL.

- **Sequential Monomethod Multistrand Designs.** In Table 2, these are designated as QUAN→QUAN or QUAL→QUAL. In these designs, there are multiple strands of the research design which occur in a sequential or chronological order. All stages within the strands (conceptualization, experiential, inferential) are consistently either QUAN or QUAL. The second (or subsequent) strand(s) of the study emerge from the outcome and inferences of a previous strand.

The Concurrent Monomethod Multistrand Design is illustrated in Figure 2. The designs in Cell Two were foreshadowed by the multimethod-multitrait matrix of Campbell and Fiske (1959). Their study presented one of the first explicit “multimethod” designs in the social and behavioral sciences. Specifically, it utilized more than one QUAN method (e.g., a structured interview that yielded QUAN data, a structured observation protocol that also yielded QUAN data) to measure a single psychological trait.<sup>6</sup> (Refer to Figure 2 for an illustration of Concurrent Monomethod multistrand Designs.)

*Mixed Methods Designs*

There are two types of MM designs, those with only one strand (Cell Three in Table 2) and those with more than one strand (Cell Four in Table 2). Cell Three designs are Mixed Methods Monostrand Designs, while Cell Four designs are Mixed Methods Multistrand designs.

*Mixed Methods Monostrand Designs.* These are the simplest of the MM designs, involving only one strand of a research study, yet including both QUAL and QUAN components. Because only one type of data is analyzed and only one type of inference (QUAL or QUAN) is made, we labeled these designs as *quasi-mixed* above.

We discuss only one design from Cell Three in detail in this section: the Monostrand Conversion Design. (Refer to Figure 3 for an illustration of Monostrand Conversion Designs.) There are some important points regarding the Monostrand Conversion Design:

- In general, conversion designs allow for data transformation where one data form is converted into the other and then analyzed accordingly. Conversion designs represent the third distinct way to implement MM designs, in addition to concurrent and sequential designs.
- **Monostrand Conversion Designs** (also known as the Simple Conversion Design) are utilized in single strand studies in which research questions are answered through an analysis of transformed data (i.e., quantitized or qualitized data). These studies are mixed because they switch approach in the experiential phase of the

study, when the data that were originally collected (narrative, numeric) are converted into the other form (numeric, narrative). Figure 3 depicts the monostrand conversion designs, with the transformation of data type occurring between the methodological and the analytical components of the experiential stage.

- Monostrand conversion designs may be planned before the study actually occurs, but many applications of this design occur serendipitously as a study unfolds. For instance, a researcher may determine that there are emerging patterns in the information gleaned from narrative interview data that can be converted into numerical form and then analyzed statistically, thereby allowing for a more thorough analysis of the data. Monostrand conversion designs are often serendipitously occurring, unplanned, emerging designs that may be employed together with other preplanned research strands.

An interesting attribute of the Monostrand Conversion Design is that it has been used extensively in both the QUAN and QUAL traditions, without being recognized as “mixed” (for examples, see Hunter & Brewer, 2003; Maxwell & Loomis, 2003; Waszak & Sines, 2003). Some of the explicit descriptions of quantizing data in the MM research literature include:

- Morse’s (1989) study of teenage mothers and the frequency of their use of the word “stuff”,

converting that word into a frequency count that demonstrated the childish mode of speech used by young women with adult responsibilities.

- Miles and Huberman’s (1994) conversion of narrative data from their school improvement studies into frequency counts or rating scales; for example, their conversion of the described “roughness” or “smoothness” of the implementation process into three-to-five point scales.
- Sandelowski, Harris, and Holditch-Davis’ (1991) transformation of interview data into a frequency distribution that compared the “numbers of couples having and not having an amniocentesis with the number of physicians encouraging or not encouraging them to have the procedure” which was then analyzed statistically to determine the “relationship between physician encouragement and couple decision to have an amniocentesis” (Sandelowski, 2003, p. 327).

*Mixed Methods Multistrand Designs.* Cell Four of Table 2 contains the Mixed Methods Multistrand Designs, which are the most complex of the designs in the matrix. All of these designs contain mixed methods and at least two research strands. Mixing of the QUAL and QUAN approaches may occur both within and across all three stages of the study. There are four of these designs, which we consider to be the most valuable

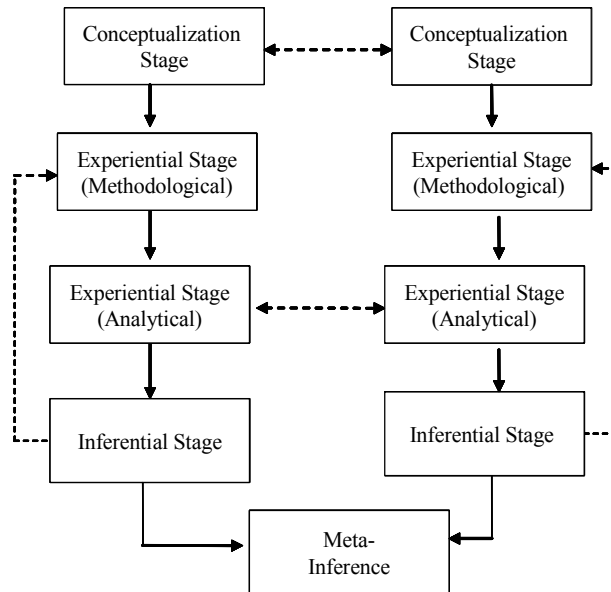


Figure 2. Graphic Presentation of Concurrent Monomethod Multistrand Designs (two QUAN strands or two QUAL strands)

of the MM designs presented in Table 2<sup>7</sup>:

- Concurrent Mixed Design
- Sequential Mixed Designs
- Conversion Mixed Designs
- Fully Integrated Mixed Designs

These four types of designs are families of designs. There may be several permutations of members of these families based on other design criteria (e.g., priority of methodological approach).

Concurrent Mixed Designs are designs in which there are at least two relatively independent strands: one with QUAL questions and data collection and analysis techniques and the other with QUAN questions and data collection and analysis techniques. Inferences made on the basis of the results from each strand are synthesized to form meta-inferences at the end of the study. (Refer to Figure 4 for an illustration of Concurrent Mixed Designs.)

A major advantage of MM research is that it enables researchers simultaneously to ask confirmatory and exploratory questions, and therefore verify and generate theory in the same study. Concurrent Mixed Designs use QUAL and QUAN data and analyses in independent strands to answer exploratory (typically, but not always, QUAL) and confirmatory (typically, but not always, QUAN) questions.

An example of such a design is a World Bank Guatemala Poverty Assessment (described in Rao & Woolcock, 2003). The QUAN strand of the study included survey data. Based on these (survey) data, a purposive sample of five pairs of villages was selected for QUAL study. The two strands of the study were kept independent (including the investigator teams) until after all data analyses were finished. Mixing happened at the inference stage. The integration provided:

... a more accurate map of the spatial and demographic diversity of the poor, as well as, crucially, a sense of the immediate context within which poverty was experienced by different ethnic groups, details of the local mechanisms that excluded them from participation in mainstream economic and civic activities, and the nature of the barriers they encountered in their efforts to advance their interests and aspirations. (Rao & Woolcock, 2003, p. 173)

Lopez and Tashakkori (2006) provide another example of a concurrent mixed study of the effects of two types of bilingual education programs on attitudes and academic achievement of fifth-grade students. The QUAN strand of the study included standardized achievement tests in various academic subjects, as well as linguistic competence in English and Spanish, and utilized a Likert-type scale measuring self-perceptions as well as self-beliefs in relation to bilingualism. The QUAL strand consisted of interviews with a random sample of

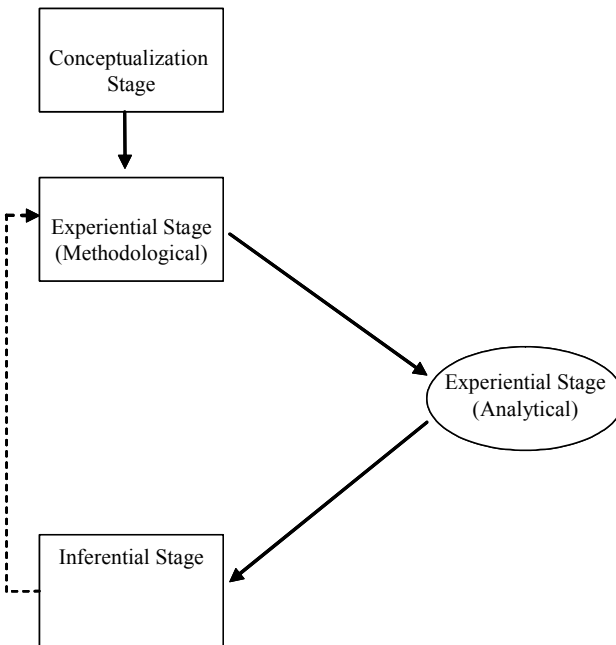


Figure 3. Graphic Presentation of Monostrand Conversion Design (Quasi-Mixed design)

32 students in the two programs. Each set of data were analyzed independently, and conclusions were drawn. The findings of the two studies were integrated by (a) comparing and contrasting the conclusions, and (b) by trying to construct a more comprehensive understanding of how the two programs impacted the children.

While Concurrent Mixed Designs are very powerful, they are challenging to conduct due to the complexity of running multiple strands of research simultaneously. As in the World Bank study mentioned above, different teams of researchers may be required to conduct these studies as was the case with a study described by Trend (1978). This study involved the concurrent, but separate, collection of QUAN and QUAL data on a federal housing-subsidy program. In this study, which was set up as a Concurrent Mixed Design<sup>8</sup>, a team of quantitatively oriented evaluators conducted the QUAN strand, whereas anthropologists conducted the QUAL strand. The components of the study were as follows:

1. The QUAN strand was set up to determine if the use of direct cash housing allowance payments would assist low-income families to obtain decent housing on the open market. The QUAN strand involved an analysis of three “pre-experiments” (Campbell & Stanley, 1963) that generated mostly survey data on agency activities, expenses, demographic characteristics of clients, and housing quality.
2. The QUAL strand involved the generation of case studies by observers using field observations, interviews, and documents. The purpose of the case studies was to provide a holistic description of what actually occurred at the program sites.

The QUAN data were expected to determine the success of the program, while the QUAL case studies were used to provide a picture of program process. Meta-inferences across the two independent strands were employed to reconcile the information gleaned from the two concurrent strands.

Concurrent Mixed Designs are difficult for novice researchers or researchers working alone to conduct for several reasons:

- In general, it requires considerable expertise to examine simultaneously and separately the same phenomenon using two different approaches.
- Specifically, the simultaneous analysis of QUAN and QUAL data sources and then the integration of those results into a coherent set of findings and inferences is difficult.
- Particular problems may develop when the results are discrepant, and the novice and/or solo investigator may be unable to interpret and/or resolve these inconsistencies in order to make meta-inferences.

Hence, the very powerful Concurrent Mixed Designs are best accomplished using a collaborative team approach in which each member of the group can contribute to the complex, simultaneously evolving research design.

Sequential Mixed Designs are designs in which there are at least two strands that occur chronologically (QUAN→QUAL or QUAL→QUAN). The conclusions that are made on the basis of the results of the first strand lead to formulation of questions, data collection, and data analysis for the next strand. The final inferences are

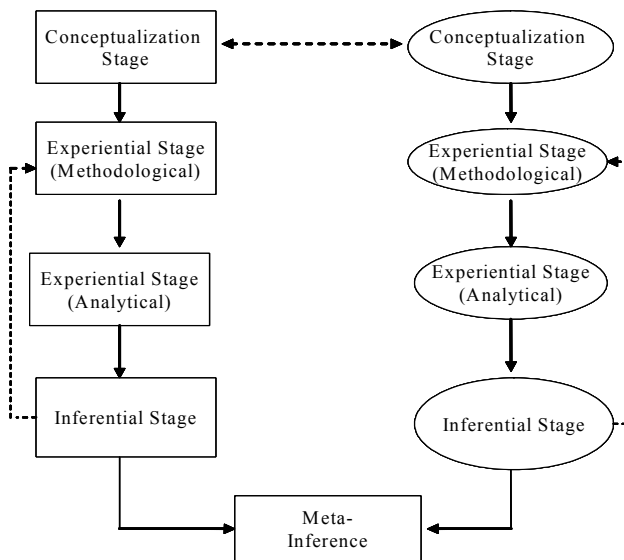


Figure 4. Graphic Presentation of Concurrent Mixed Design

based on the results of both strands of the study. The second strand of the study is conducted either to confirm/disconfirm the inferences of the first strand or to provide further explanation for findings from the first strand (Tashakkori & Teddlie, 2003, p. 715). (Refer to Figure 5 for an illustration of Sequential Mixed Designs.)

Sequential Mixed Designs answer exploratory and confirmatory questions chronologically in a pre-specified order. While still difficult, these designs are easier to conduct by the solo investigator than the Concurrent Mixed Designs, because it is easier to keep the strands separate and the studies typically unfold slower and in a more predictable manner.

We have had a series of graduate students conduct dissertations with Sequential Mixed Designs over the past several years (e.g., Aaron, 2005; Cakan, 1999; Kochan, 1998; Lasserre-Cortez, 2006; Stevens, 2001; Wu, 2005). Some of these dissertations used more complex combinations of the two approaches rather than a simple sequential study. For example, Wu's (2005) dissertation consisted of a QUAL strand (administrator interviews) as well as a quantitative (survey) one. In each strand, data were collected and analyzed independently, similar to a concurrent mixed design. However, the strongest inferences were gleaned when a sequential data analysis was performed in which the themes obtained from the QUAL strand were used for comparison of the QUAN data. The inconsistency between the inferences of the two strands was the most striking conclusion from the study, in that it revealed a gap between the student applicants' and college

administrators' perceptions of factors impacting college choice in Taiwan.

It is often difficult for students to think both retrospectively and prospectively when developing and defending a proposal, especially when they are just beginning to master the knowledge base in a given area. Students like the Sequential Mixed Designs because they allow them to address some issues already discussed in the literature in one phase of the study and then contribute to the knowledge base in an exploratory manner in another phase of the study.

An example of a Sequential QUAL→QUAN Mixed Design comes from the consumer marketing literature (Hausman, 2000). The first part of the study was exploratory in nature utilizing semi-structured interviews to examine several questions related to impulse buying. Results from these interviews were then used to generate a series of hypotheses related to this phenomenon. The semi-structured interviews in the first part of the study examined several research questions:

- What are consumers' attitudes toward shopping?
- How do consumers make buying decisions?
- How do buying decisions result in impulse buying?

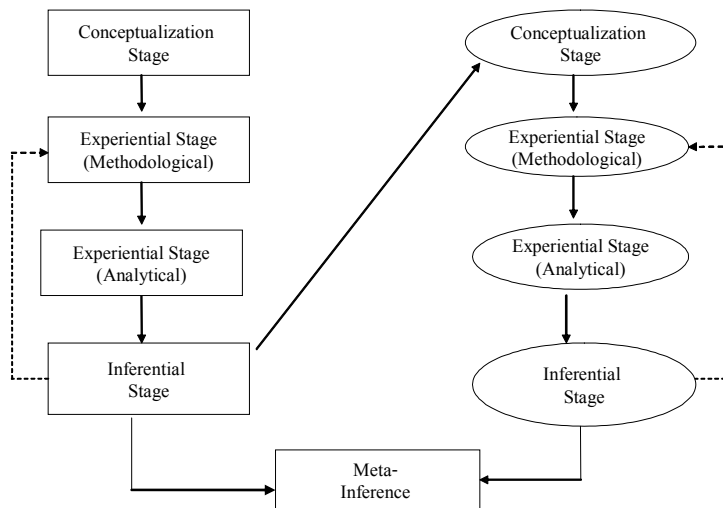


Figure 5. Graphic Presentation of Sequential Mixed Design

Trained interviewers conducted 60 interviews with consumers, and the resultant data were analyzed using grounded theory techniques. Hausman (2000, p. 406) noted that this “methodology yields hypotheses empirically grounded in the data.”

Based on these analyses, a series of five hypotheses were developed and tested using a 75-item questionnaire generated for the purposes of this study. A final sample of 272 consumers completed the questionnaire. Hypothesis testing involved both correlational and analysis of variance techniques. Significant results were identified for three of the hypotheses:

- Individual consumers’ impulse buying is correlated with their desires to fulfill hedonic needs (e.g., fun, novelty, surprise).
- Individual consumer impulse buying behavior is correlated with desires to satisfy self-esteem considerations.
- Perceptions of decision-making accuracy mediate impulse buying.

The Conversion Mixed Design is a multistrand concurrent design in which mixing of QUAL and QUAN approaches occurs in all components/stages, with data transformed (qualitized or quantitized) and analyzed both qualitatively and quantitatively (Tashakkori & Teddlie, 2003, p. 706). In these designs, one type of data (e.g., QUAL) is gathered and is analyzed accordingly (QUAL) and then transformed and analyzed using the other methodological approach (e.g., quantitized).

Witcher, Onwuegbuzie, Collins, Filer, and Wiedmaier (2003) conducted a Conversion Mixed Design, which was subsequently described by Onwuegbuzie and Leech (2004). In this study, the researchers gathered QUAL data from 912 undergraduate and graduate students regarding their perceptions of the characteristics of effective college teachers. A QUAL thematic analysis revealed nine characteristics of effective college teachers, including student-centeredness and enthusiasm about teaching. The researchers then quantitized the data by assigning binary values for each of the students for each of the themes. Thus, if a female graduate student made responses that indicated that she thought student-centeredness was a characteristic of effective college teaching, that student received a score of “1” for that theme. On the other hand, if another student did not make responses indicating that he thought student-centeredness was a characteristic of effective college students, that student received a score of “0” for that theme. A series of binary codes (1, 0) were assigned to each student for each characteristic of effective teaching, resulting in what Witcher et al. (2003) called an inter-respondent-matrix (participant X theme matrix).

The analysts then subjected these quantitized (“binarized”) data to a series of analyses that enabled

them to statistically associate each of the nine themes of effective college teaching with four demographic variables (gender, race, undergraduate or graduate status, and preservice service status). The researchers were able to connect students with certain demographic characteristics with preferences for certain effective teaching characteristics (e.g., females were more likely to endorse student-centeredness than were male students).

Thus, one type of data (QUAL) was subjected to both thematic and statistical analysis and meta-inferences were possible using both types of data simultaneously. Onwuegbuzie and Leech (2004) concluded that “... subjecting quantitized data to statistical analysis aided Witcher et al. in the interpretation of the qualitative themes” (p. 784).

The Fully Integrated Mixed Design, to use a British phrase, is the “Full Monty” of MM designs: a multistrand concurrent design in which mixing of QUAL and QUAN approaches occurs in an interactive (i.e., dynamic, reciprocal, interdependent, iterative) manner at all stages of the study. At each stage, one approach (e.g., QUAL) affects the formulation of the other (e.g., QUAN) (Tashakkori & Teddlie, 2003, p. 708). (Refer to Figure 6 for an illustration of Fully Integrated Mixed Design.)

The Louisiana School Effectiveness Study (Teddlie & Stringfield, 1993) included a longitudinal study of eight matched pairs of schools initially classified as either effective or ineffective using baseline achievement score data collected at time one (T<sub>1</sub>, 1982-84) across two phases of the study (T<sub>2</sub>, 1984-85 and T<sub>3</sub>, 1989-90). There were two basic questions that characterized the longitudinal third and fourth phases of the study (LSES-III and IV with data gathering conducted at T<sub>2</sub> and T<sub>3</sub> respectively):

- Would the eight matched pairs of more effective and less effective schools remain differentially effective over time, or would some schools increase or decrease in effectiveness status over time? The major QUAN data used to answer this question were achievement scores and indices of student socioeconomic status.
- What are the processes whereby schools remain the same or change over time with regard to how well they educate their students? The major QUAL data used to answer this question were classroom and school level observations and interviews with students, teachers, and principals.

The LSES employed a Fully Integrated Mixed Design as follows:

- At the conceptualization stage, the formulation of the QUAN oriented questions informed the

formulation of the QUAL oriented questions, and vice versa.

- At the experiential (methodological/analytical) stage, QUAL data were quantitized and analyzed statistically, and QUAN data were qualitized and profiles of schools were generated. The results of these statistical and profile analyses further affected the formulation of additional QUAL and QUAN analyses.
- The two major QUAN and QUAL strands, and their crossover analyses, directly influenced the formulation of the meta-inferences, which resulted in a dozen or so major conclusions, each of which involved triangulated data.
- This complex design was accomplished with the services of a research team that had a wide variety of methodological and experiential backgrounds, as recommended by Shulha and Wilson (2003).

As noted throughout this article, there are also quasi-mixed multi-strand designs in Cell Four of Table 2. It should be evident to the reader that in the multi-strand designs, one approach/strand might only be a small part of the overall study. For example, in a Concurrent Quasi-Mixed Study, limited QUAN survey data might be collected and analyzed, in order to provide insights about a relatively larger group of respondents than the QUAL study was able to generate.

Examples of studies in which one or the other approach minimally contributes to the final inferences are the following:

- A sociological study conducted in the QUAN tradition with hypotheses predicting significant relationships between several predictor variables (e.g., self report items measuring length of marriage, number of children, feelings of affection toward spouse, age, gender) and marital satisfaction (another self report item). In addition to the QUAN self report items, the participants were asked to complete an open-ended item asking them to define what “marital satisfaction” meant to them. The most important data were the participants’ responses to the QUAN items which were analyzed statistically to test the predicted relationships, but the complementary QUAL information on what the couples thought “marital satisfaction” meant provided interesting “side bar” results. The inferences from the study were made deductively within the postpositivist framework of sociological prediction studies, but the conclusions also included some anecdotal evidence from the participants.
- An ethnographic study of a large city police force conducted by an anthropologist who had gained entry into the social setting and was operating as a participant observer. The orientation of the study was inductive and constructivist in nature, and the most important data were the observations and interviews that the anthropologist conducted. The researcher also collected some QUAN information in the

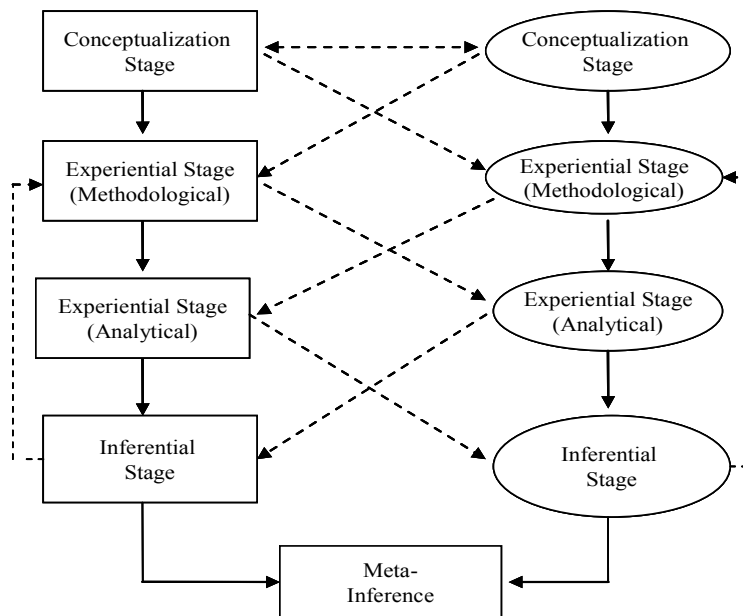


Figure 6. Graphic Presentation of Fully Integrated Mixed Design

form of work time data (number of hours spent in various activities during shifts), percentage of ethnic minorities and women on the police force, percentage of minorities and women among those arrested, and so forth. The inferences from the study were made within the constructivist framework of the ethnographic anthropology, but the conclusions also included interesting QUAN data that helped strengthen some interpretations.

#### Summary

This article presents the Methods-Strands Matrix, which features MM research designs, especially four families of MM designs: sequential, concurrent, conversion, and fully integrated. Before discussing the matrix, we briefly discussed (a) why typologies of MM research designs are useful and (b) the dimensions that have been used by various authors to create MM design typologies.

We devoted much of the article to an extended discussion of our typology of MM research designs illustrated in the Methods-Strands Matrix. This matrix conceptually includes all designs, but emphasizes the mixed methods ones, which were presented as a family of designs that differ with regard to three key criteria and one less important criterion:

- Number of methodological approaches used
- Number of strands in the research design
- Type of implementation process
- Stage of integration - a distinction was made between Mixed Methods Designs (in which integration happens in more than one stage) and Quasi-Mixed Designs (in which there are two types of data and analysis, but no integration of the inferences occurs).

Four families of mixed methods designs were featured in the matrix. Examples of these families of designs were presented throughout the article.

Based on the information presented in this article and other sources, we (Teddlie & Tashakkori, forthcoming) have developed a seven step process for researchers selecting the best design for their projects from our matrix or another of the available typologies:

1. The researcher must first determine if her research questions require a monomethod or MM design.
2. The researcher should be aware that there are a number of typologies of MM research designs and should know how to access details regarding them.

3. The researcher wants to select the best MM research design for her particular study and assumes that one of the published typologies includes the right design for her project.
4. Typologies may be differentiated by the criteria that are used to distinguish among the research designs within them, and the researcher needs to know those criteria.
5. These criteria should be listed by the researcher, who may then select the criteria that are most important to her for the particular study she is designing.
6. The researcher then applies the selected criteria to potential designs, ultimately selecting the best research design for her study.
7. In some cases, the researcher may have to develop a new mixed methods design, because no one best design exists for her research project.

The lead editors for this article were R. Burke Johnson and Anthony J. Onwuegbuzie.

#### References

- Aaron, L. S. (2005). *Responsibilities and leadership styles of radiologic technology program directors: Implications for leadership development*. Unpublished doctoral dissertation, Louisiana State University, Baton Rouge.
- Cakan, M. (1999). *Interaction of cognitive style and assessment approach in determining student performance on tests of second language proficiency*. Unpublished doctoral dissertation, Louisiana State University, Baton Rouge.
- Campbell, D., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Campbell, D. T., & Stanley, J. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 171-246). Chicago: Rand McNally. (Also published as *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally, 1966.)
- Cook, T. D., & Campbell, D. T. (1979). *Quasiexperimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin Company.



- Creswell, J., Plano-Clark, V., Gutmann, M., & Hanson, W. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209-240). Thousand Oaks, CA: Sage.
- Fetterman, D. M. (1998). *Ethnography: Step by step* (2nd ed.). Thousand Oaks, CA: Sage.
- Greene, J., & Caracelli, V. (Eds.). (1997). *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms* (New Directions for Evaluation, No. 74). San Francisco: Jossey-Bass.
- 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.
- Hausman, A. (2000). A multi-method investigation of consumer motivations in impulse buying behavior. *Journal of Consumer Marketing, 17*, 403-419.
- Hunter, A., & Brewer, J. (2003). Multimethod research in sociology. In A. Tashakkori, and C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 577-594). Thousand Oaks, CA: Sage.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). *Mixed methods research: A research paradigm whose time has come. Educational Researcher, 33*(7), 14-26.
- Kochan, S. (1998). *Considering outcomes beyond achievement: Participation as an indicator of high school performance*. Unpublished doctoral dissertation, Louisiana State University, Baton Rouge.
- Lasserre-Cortez, S. (2006). *A mixed methods examination of professional development through whole faculty study groups*. Unpublished doctoral dissertation, Louisiana State University, Baton Rouge.
- LeCompte, M. D., & Preissle, J. (1993). *Ethnography and qualitative design in educational research* (2nd ed.). New York: Academic Press.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills: Sage.
- Lopez, M. & Tashakkori, A. (2006). Differential outcomes of TWBE and TBE on ELLs at different entry levels. *Bilingual Research Journal, 30*(1), 81-103.
- Maxwell, J., & Loomis, D. (2003). Mixed methods design: An alternative approach. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 241-272). Thousand Oaks, CA: Sage.
- Mertens, D. M. (2005). *Research methods in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Miles, M., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook*. (2nd ed.) Thousand Oaks, CA: Sage.
- Morgan, D. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative Health Research, 8*, 362-376.
- Morse, J. (1989). *Qualitative nursing research: A contemporary dialogue*. Newbury Park, CA: Sage.
- Morse, J. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research, 40*, 120-123.
- Morse, J. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 189-208). Thousand Oaks, CA: Sage.
- Onwuegbuzie, A. J., & Leech, N. L. (2004). Enhancing the interpretation of “significant” findings: The role of mixed methods research. *The Qualitative Report, 9*, 770-792. Retrieved March 6, 2006, from <http://www.nova.edu/ssss/QR/QR9-4/onwuegbuzie.pdf>
- Patton, M. Q. (1990). *Qualitative research and evaluation methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Rao, V., & Woolcock, M. (2003). Integrating qualitative and quantitative approaches in program evaluation. In F. J. Bourguignon & L. Pereira da Silva (Eds.), *Evaluating the poverty and distribution impact of economic policies* (pp. 165-190). New York, N.Y.: Oxford University Press.
- Sandelowski, M. (2003). Tables or tableaux? The challenges of writing and reading mixed methods studies. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 321-350). Thousand Oaks, CA: Sage.
- Sandelowski, M., Harris, B. G., & Holditch-Davis, D. (1991). Amniocentesis in the context of infertility. *Health Care for Women International, 12*, 167-178.
- Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for general causal inference*. Boston: Houghton Mifflin.

- Shulha, L. M., & Wilson, R. J. (2003). Collaborative mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 639-670). Thousand Oaks, CA: Sage.
- Stevens (2001). *Differential modes of external change agent support in diffusion of innovation*. Unpublished doctoral dissertation, Louisiana State University, Baton Rouge.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining the qualitative and quantitative approaches*. Thousand Oaks, CA: 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 and behavioral research* (pp. 671-702). Thousand Oaks, CA: Sage.
- Teddlie, C., & Stringfield, S. (1993). *Schools make a difference: Lessons learned from a 10-year study of school effects*. New York: Teachers College Press.
- Teddlie, C., & Tashakkori, A. (2005, April). *The methods-strands matrix: A general typology of research designs featuring mixed methods*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.
- Teddlie, C., & Tashakkori, A. (in press). *Foundations of mixed methods research: Integrating quantitative and qualitative techniques in the social and behavioral sciences*. Thousand Oaks, CA: Sage.
- Trend, M. G. (1978). On the reconciliation of qualitative and quantitative analyses: A case study. *Human Organization*, 37, 345-354. (Reprinted in T. D. Cook & C. S. Reichardt (Eds.), *Qualitative and quantitative methods in evaluation research*. Thousand Oaks, CA: Sage.)
- Waszak, C., & Sines, M. (2003). Mixed methods in psychological research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 557-576). Thousand Oaks, CA: Sage.
- Witcher, A. E., Onwuegbuzie, A. J., Collins, K. M. T., Filer, J., & Wiedmaier, C. (2003, November). *Students' perceptions of characteristics of effective college teachers*. Paper presented at the annual meeting of the Mid-South Educational Research Association, Biloxi, MS.
- Wu, C. (2005). *Correlates of college choice among Taiwanese youth: Relative importance of personal, social, and institutional considerations*. Unpublished doctoral dissertation, Florida International University, Miami.

#### Notes

<sup>1</sup>This article is based on a paper presented by the authors at the 2005 Annual Meeting of the American Educational Research Association (Teddlie & Tashakkori, 2005).

<sup>2</sup>Our typology actually emphasizes the first three criteria (number of methodological approaches, number of strands, type of implementation process) as described throughout this text. The fourth criterion, stage of implementation, is retained because it allows for the inclusion of quasi-mixed designs, which are defined later in this article.

<sup>3</sup>Therefore, we have dropped the term “model” to avoid confusion. Our previously defined (Tashakkori & Teddlie, 2003) “mixed model designs” are now simply “mixed designs”. Our previously defined “mixed designs” are now “quasi-mixed designs.”

<sup>4</sup>See Tashakkori and Teddlie (2003, pp. 685-689) for more details regarding these designs (mixed at the experiential stage only).

<sup>5</sup>It is important to note that multistrand designs are illustrated as having only two strands throughout this chapter for the sake of simplicity. They could be more complex, involving three or more strands (e.g., QUAL→QUAN→QUAL).

<sup>6</sup>What Campbell and Fiske (1959) called “multimethod” is what we call “multistrand” in Cell Two with two exceptions: (1) they referred only to QUAN designs, while we also refer to QUAL methods and (2) their model emphasized the methods stage alone, while we have added the conceptualization and inferential stages in our matrix.

<sup>7</sup>Once again, if mixing occurs in the experiential stage only, then these designs should be considered *quasi-mixed*.

<sup>8</sup>Throughout this chapter, we refer to particular studies as being examples of designs from our matrix. The authors of the original studies did not use these design names, because they were not introduced into the research literature until 2003. Our designation of these

studies as particular types of mixed methods designs is based on an ex post facto analysis of their characteristics.

## Defining and Designing Mixed Research Synthesis Studies

**Margarete Sandelowski**

*University of North Carolina at Chapel Hill School of Nursing*

**Corrine I. Voils**

*Durham Veterans Affairs Medical Center  
Duke University Medical Center Department of Medicine*

**Julie Barroso**

*Duke University School of Nursing*

*Mixed research synthesis is the latest addition to the repertoires of mixed methods research and systematic review. Mixed research synthesis requires that the problems generated by the methodological diversity within and between qualitative and quantitative studies be resolved. Three basic research designs accommodate this diversity, including the segregated, integrated, and contingent designs. Much work remains to be done before mixed research synthesis can secure its place in the repertoires of mixed methods research and systematic review, but the effort is well worth it as it has the potential to enhance both the significance and utility for practice of the many qualitative and quantitative studies constituting shared domains of research.*

Mixed research synthesis is the latest addition to the repertoires of mixed methods research and systematic review. *Mixed research synthesis* is our name for the type of systematic review aimed at the integration of results from both qualitative and quantitative studies in a shared domain of empirical research. In contrast to mixed methods research in which the data set subject to analysis and interpretation is composed of the qualitative and quantitative *data* (e.g., from interviews, observations, questionnaires, physiologic measures, and the like) obtained directly from research participants within a single study or program of research, the data in mixed research synthesis studies are the *findings* of primary qualitative and quantitative studies in a designated body of empirical research. The focus of mixed research synthesis studies is on researchers' integrations of their data, or the results they report; the products of mixed research synthesis studies are other researchers' (i.e., reviewers of research) integrations of those results to "sum up" what is known about a target phenomenon and, thereby, to direct both practice and future research.

In this article, we offer an overview of the impetus for mixed research synthesis and the challenges it presents, and propose three basic research designs to conduct mixed research synthesis studies. We have drawn from the large body of literature in the general areas of systematic review, research synthesis, and mixed methods research; the small body of literature on mixed research synthesis; and from our own on-going study aimed at developing methods to synthesize qualitative and quantitative research findings.<sup>1</sup>

### The Impetus for Mixed Research Synthesis

The new interest in mixed research synthesis (Dixon-Woods, Agarwal, Young, Jones, & Sutton, 2004; Forbes & Griffiths, 2002; Harden & Thomas, 2005; Hawker, Payne, Kerr, Hardey, & Powell, 2002; Lemmer, Grellier, & Steven, 1999; Mays, Pope, & Popay, 2005; Popay & Roen, 2003) is the result of the convergence of two "growth industries" (Estabrooks, 1999, p. 274) discussed in the following sections, namely, evidence-based practice and qualitative research.

### *The Turn to Evidence-Based Practice*

Over the last two decades, scholars in the practice disciplines have increasingly turned to evidence-based practice to facilitate better use of research findings and to close the research-practice gap. Appearing in various

---

Correspondence for this article should be addressed to Margarete Sandelowski, 7460 Carrington Hall, Chapel Hill, NC 27599.  
Email: msandelo@email.unc.edu.

guises across the disciplines as evidence-based health care, medicine, nursing, education, social work, and librarianship (Trinder & Reynolds, 2000), evidence-based practice is generally defined as the conscientious, explicit, and judicious use of information to serve as the foundation for practice (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996; Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000). At the heart of evidence-based practice are evidence syntheses, or integrations of research findings derived from systematic reviews of empirical research in targeted research areas to answer specific research questions addressing specific practice problems. Such evidence syntheses are viewed as having the potential to increase the utility of research and the effectiveness of practice as they enable answers to such questions as which treatments of a disease produce the best health outcomes, or which teaching strategies produce the best learning outcomes.<sup>2</sup>

### *The Rise of Qualitative Research*

Concurrent with the turn to evidence-based practice has been the growth of qualitative research. Over the last 30 years, the number of qualitative studies and of instructional texts on qualitative methods has increased exponentially across the behavioral, social science, and practice disciplines. The dramatic proliferation of qualitative studies and rising concern about their under-utilization, occurring against the backdrop of renewed interest in enhancing the utility of research through systematic reviews of research, sparked the interest in conducting syntheses of qualitative research (Sandelowski, 2004). Most commonly referred to as qualitative metasynthesis or meta-ethnography, qualitative research synthesis studies were promoted to fulfill the promise of qualitative research findings to effect desired changes in health, education, and social welfare. A spate of articles and books has appeared since the late 1980s addressing qualitative research synthesis methods (e.g., Noblit & Hare, 1988; Paterson, Thorne, Canam, & Jillings, 2001; Sandelowski & Barroso, in press), reporting the results of qualitative research synthesis studies (e.g., Campbell et al., 2003; Pound et al., 2005; Sandelowski & Barroso, 2003, 2005), and calling for the inclusion of qualitative research into evidence-based practice (Barbour, 2000; Dixon-Woods, Fitzpatrick, & Roberts, 2001; Giacomini, 2001; Green & Britten, 1998; Greenhalgh, 2002; Leys, 2003; Popay & Williams, 1998).

### The Challenges of Mixed Research Synthesis

Advances in qualitative and quantitative research synthesis and the increasing prominence of mixed methods research as the “third research paradigm”

(Johnson & Onwuegbuzie, 2004, p. 14) for the practice disciplines have contributed to the current interest in mixed research synthesis. But for mixed research synthesis to advance, researchers must solve the problems generated by the methodological diversity within and between qualitative and quantitative studies. Difference has recurrently been identified as the most important factor complicating both the qualitative and quantitative research synthesis enterprises (Cooper, 1998; Mulrow, Langhorne, Grimshaw, 1997; Sandelowski, Docherty, & Emden, 1997; Sandelowski, Voils, & Barroso, 2006). Even studies of ostensibly the same variables or target events, experiences, or phenomena in ostensibly similar groups of people employing ostensibly the same methodological approaches will have differences sufficient to require finding the means to enable their findings to be compared and combined.

### *Qualitative Research Synthesis*

In the qualitative research synthesis literature, the difference problem is most often addressed in relation to the philosophical differences among research traditions and the singularity of every research participant and research encounter (Sandelowski et al., 1997). The qualitative research methods literature is characterized, in part, by efforts to differentiate among ontological positions (e.g., realist, idealist, and relativist); epistemological positions (e.g., objectivist, constructionist, subjectivist); paradigms of inquiry (e.g., neo-positivism, interpretivism, critical theory, postmodernism); foundational theories and philosophies (e.g., symbolic interactionism, Heideggerian phenomenology, Foucaultian genealogy); and methodologies (e.g., grounded theory, phenomenology, ethnography, narrative/discourse study; Crotty, 1998; Guba & Lincoln, 2005).<sup>3</sup> These differences can make the synthesis of qualitative findings alone a daunting enterprise. In addition, the qualitative research emphasis on delineating the complexities and contradictions of “N=1 experiences” (Eisner, 1991, p. 197) seems at odds with the conventional research synthesis emphasis on simplification and summary. Moreover, the diversity in the implementation and reporting of qualitative research complicates both the identification of the methods actually used in a study and the findings produced from those methods (Sandelowski & Barroso, in press).

Accordingly, qualitative researchers have urged the development of synthesis methods distinctive to qualitative inquiry and warned against reliance on quantitative research synthesis as the model for qualitative research synthesis (Barbour & Barbour, 2003; Jones, 2004; Sandelowski et al., 1997). Although much progress has been made in the development of

these methods, debates continue over such issues as: (a) terminology (e.g., qualitative meta-analysis [or metaanalysis], meta-synthesis [or metasynthesis], meta-data-analysis); (b) whether to retrieve all of the research reports in a targeted domain or only a purposeful sample of them; (c) whether and how to use quality criteria for evaluating qualitative studies; (d) what the goals of qualitative research synthesis should be (e.g., topical review, aggregation, integration, interpretive comparison, critique); and (d) whether to advance one or multiple interpretations of studies (Booth, 2001; Doyle, 2003; Evans & Pearson, 2001; Jones, 2004; Thorne, Jensen, Kearney, Noblit, & Sandelowski, 2004). Because of the “multidisciplinary pedigree” (Barbour & Barbour, 2003, p. 183) of qualitative research, the resolution of these debates will vary with individual research practitioners’ understanding of the imperatives of qualitative research as applied to their own disciplines.

#### *Quantitative Research Synthesis*

In the quantitative research literature, the problem of difference is referred to as heterogeneity and is addressed as the methodological diversity within and between observational and experimental studies, and in terms of the contrast between the “real” differences in the target phenomena and “artifactual” differences owing to the way these phenomena were studied (Glasziou & Sanders, 2002). The constellation of statistical techniques known as meta-analysis, which is most identified with quantitative research synthesis and intended to accommodate these differences, continues to engender debate. One continuing criticism is that, because no two studies can ever be perfectly identical, meta-analysts are comparing apples and oranges (Glass, 2000; Hunt, 1997), thereby, calling into question the validity and generalizability of meta-analyses (Matt, 2003). Another criticism is that meta-analysis involves testing hypotheses about parameters in populations of studies, yet the criteria for inferential statistics (e.g., random sampling) are rarely met. Yet other concerns include the management of variable primary study quality (Conn & Rantz, 2003) and of results from diverse research designs (e.g., combining findings from single group pre-post designs and randomized controlled trials, from cross-sectional and longitudinal studies, or from experimental and observational studies). The Cochrane Collaboration and Library, an icon of the evidence-based practice movement in health care, includes only randomized controlled trials in their research syntheses, but most studies conducted that could serve as the basis for clinical practice are not randomized controlled trials.

Although many strategies have been proposed to address these issues, no consensus exists on any one set of strategies. For example, a variety of methods exist to

address the heterogeneity among studies (Higgins & Green, 2005). In recent years, meta-analysis has moved away from approaches directed toward common estimates and fixed-effects models and toward approaches estimating the extent and sources of heterogeneity among studies and random-effects models (Stangl & Berry, 2000). In addition, a variety of strategies exist to test for bias (e.g., publication bias; Schulze, Holling, & Bohning, 2003; Stangl & Berry, 2000). In the end, the desire for greater objectivity in reviewing and drawing conclusions from studies that remains the impetus for quantitative research synthesis has been curbed by the need to make highly subjective and idiosyncratic judgments that fit the nature of the primary study findings to be integrated. Despite the drive for greater objectivity evident in the growing popularity of meta-analysis, objectivity continues to reside in the defense and documentation of largely subjective judgments.

#### *Mixed Research Synthesis*

Mixed research synthesis complicates the difference problem as qualitative and quantitative research are themselves viewed as exemplifying difference. Mixed research synthesis entails the “mixing” of the differences characterizing efforts to integrate qualitative research findings with the differences characterizing efforts to integrate quantitative research findings (Sandelowski, Voils, & Barroso, 2006). Scholars have debated whether these differences preclude mixed research synthesis.

For “purists” (Johnson & Onwuegbuzie, 2004, p. 14) who view qualitative and quantitative research as two wholly different species of inquiry, the chasm between qualitative and quantitative modes of inquiry is deep enough to make it difficult or even impossible to cross it without endangering the imperatives and integrity of one or both domains of inquiry. Hammersley (2001, p. 544) cautioned against using a “positivist model” of systematic review, appropriate when using quantitative meta-analysis, but not appropriate for synthesizing qualitative findings. Drawing from Freese’s (1980) work on cumulative knowledge, she suggested that meta-analysis derives from an “additive” orientation to knowledge development, while qualitative research synthesis may require a largely “multiplicative” one (Hammersley, 2001, p. 548). Only recently have qualitative research results been considered worthy enough even to be considered for inclusion in evidence-based practice as both *good research* and *evidence* have been defined in ways that favor highly controlled quantitative studies and, thereby, automatically exclude qualitative studies (Hampton, 2002; Madjar & Walton, 2001; McKenna, Cutcliffe, & McKenna, 1999; Miller & Fredericks, 2003; Mitchell, 1999; Ray & Mayan, 2001; Upshur,

2001). In the hierarchy of evidence in which the randomized clinical trial is the gold standard against which all studies are evaluated as good or bad research and as yielding strong or weak evidence, qualitative research will inevitably be ranked at the bottom (Evans, 2003; Petticrew & Roberts, 2003).

Proponents of qualitative research have expressed concern about the cooptation of qualitative research that occurs when its distinctive imperatives appear to be celebrated when they are actually only tolerated (Sandelowski & Barroso, in press). Rolfe (2002) described the veneer of acceptance of diversity that often masks efforts toward conformity, or the erasure of difference. Evidence of cooptation may be seen in the prevailing solutions proposed in the mixed research synthesis literature to the problems of “mixing” qualitative and quantitative research results. These solutions include primarily the one-way assimilation of qualitative data into quantitative data, or the use of qualitative data largely as an adjunct to quantitative research synthesis. Few solutions involving the assimilation of quantitative data into qualitative data have been proposed probably because any “narrative” or “qualitative” solution is viewed as reverting to the type of subjective review that the more objective quantitative meta-analysis was supposed to replace. Noteworthy here is the recurring association of the words *narrative* and *qualitative* with reviews considered unscientific and unworthy in the hierarchy of evidence in which randomized clinical trials and quantitative meta-analyses are placed on the top rungs. Conceived as useful for such purposes as clarifying the objectives of largely quantitative research reviews, delineating criteria for inclusion of studies, identifying key variables for analysis, and explaining and appraising the practical significance of reviews (Dixon-Woods, Fitzpatrick, & Roberts, 2001), qualitative research is rarely viewed as having any utility in the research synthesis enterprise outside an accessory role.

The mere toleration (as opposed to real acceptance) of qualitative research as an equal partner with quantitative research in systematic review is also apparent in the argument that because all clinical or practice problems have not been, or cannot be, addressed with the gold standard randomized clinical trial, room has to be made (albeit reluctantly) for research that is less controlled and, therefore, more biased. If this were not true, there would be no need to include these “weaker” forms of evidence. Although tolerators of qualitative research may concede that appropriately conducted study designs other than the randomized controlled trial may be stronger sources of evidence than clinical trials inappropriately conducted (i.e., that they offer the “best evidence” available [Slavin, 1995]), the randomized controlled trial remains for many the gold standard against which all

modes of inquiry are judged. Accordingly, for tolerators, qualitative research enters evidence-based practice by default, not by design.

Yet for proponents of qualitative research, it enters evidence-based practice, not only by design, but also out of necessity. Proponents of qualitative research view it as essential to achieving the goals of evidence-based practice because of its distinctive capacity for reaching facets of human experience unreachable with quantitative methods, and because of its central role in the development and testing of culturally-sensitive instruments and participant-centered interventions, and in enhancing the practical significance, and even salvaging, of quantitative research findings (e.g., Onwuegbuzie & Leech, 2004; Pope & Mays, 1995; Sandelowski, 2004; Weinholtz, Kacer, & Rocklin, 1995).

### Designing Mixed Research Synthesis Studies

Researchers’ views of the nature and impact of the differences between qualitative and quantitative research will influence how they design mixed research synthesis studies. Table 1 shows three basic designs for conducting mixed research synthesis studies that are our adaptations of designs used in primary mixed methods research (Creswell, Plano Clark, Gutmann, & Hanson, 2003; Tashakkori & Teddlie, 2003). These three designs—segregated, integrated, and contingent—accommodate different views of the relationship between qualitative and quantitative research findings and different definitions of mixed research synthesis.

#### *Segregated Design*

The segregated design shown in Table 1 maintains the conventional binary distinction between qualitative and quantitative research. This design is based on the assumptions that: (a) qualitative and quantitative studies are wholly different entities and, therefore, ought to be treated separately; (b) qualitative and quantitative studies can readily be distinguished from each other; (c) the differences between qualitative and quantitative studies warrant separate analyses and syntheses of their findings; (d) syntheses of qualitative findings require methods developed just for synthesizing qualitative findings; and (e) syntheses of quantitative findings require methods developed just for synthesizing quantitative findings. The synthesis of qualitative findings produced from such methods as qualitative metasummary, constant targeted comparison, and reciprocal translation of concepts (Sandelowski & Barroso, in press), are combined with the synthesis of quantitative findings produced from meta-analysis techniques to configure a mixed research synthesis. Only after each set of qualitative and

quantitative findings in a common domain of research has been separately synthesized with methods distinctive to it can the separate synthesis products themselves be synthesized (e.g., into a set of conclusions, a theoretical framework, or path analysis). The segregated design is most appropriate when: (a) qualitative and quantitative findings in a designated body of research are viewed as complementing (as opposed to either confirming or refuting) each other and when (b) mixed research synthesis is defined as the configuration (as opposed to the assimilation) of research findings.

*Complementarity vs. confirmation/refutation.*

Confirmation and refutation are processes that rest on the assumption that qualitative and quantitative studies can address the same research purposes or answer the same research questions and, thereby, yield findings about the same aspect of a target phenomenon. Confirmation and refutation are exercises in seeking to establish convergent validation (or triangulation) both within the qualitative and quantitative studies, respectively, and between qualitative and quantitative studies in a shared domain of research. Confirmation occurs when the same finding (e.g., that depression does or does not contribute to antiretroviral non-adherence) is repeated within and across both qualitative and quantitative studies. Refutation occurs when a designated relationship yields divergent findings, or findings in direct opposition. For example, one set of qualitative and/or quantitative studies indicates that depression contributes to non-adherence, while another set of studies addressing the same relationship indicates the opposite conclusion (e.g., that depression has no influence on adherence).<sup>4</sup>

Whereas confirmation and refutation rest on the assumption that qualitative and quantitative research can address the same research questions and, thereby, yield findings about the same aspects of phenomena, complementarity rests on the assumption that qualitative and quantitative research differ, in part, because they do not address the same questions. Barbour and Barbour (2003, p. 180) observed that qualitative research answered questions different from quantitative research as they engaged “a different sort of curiosity.” (Complicating the difference problem in mixed research synthesis and, therefore, the selection of design is that whereas the research question in quantitative studies is always fixed prior to beginning them, in qualitative studies, the research question that will ultimately be answered is often the product of analysis, or arrived at only after entering the field of study). Because they address different aspects or dimensions of a target phenomenon, qualitative and quantitative research findings can neither confirm nor refute, but rather only complement, each other. Complementarity here rests on the conception of

findings as related to each other—in that they are in the same domain of research (e.g., antiretroviral adherence)—but not as addressing the same aspects in that domain.

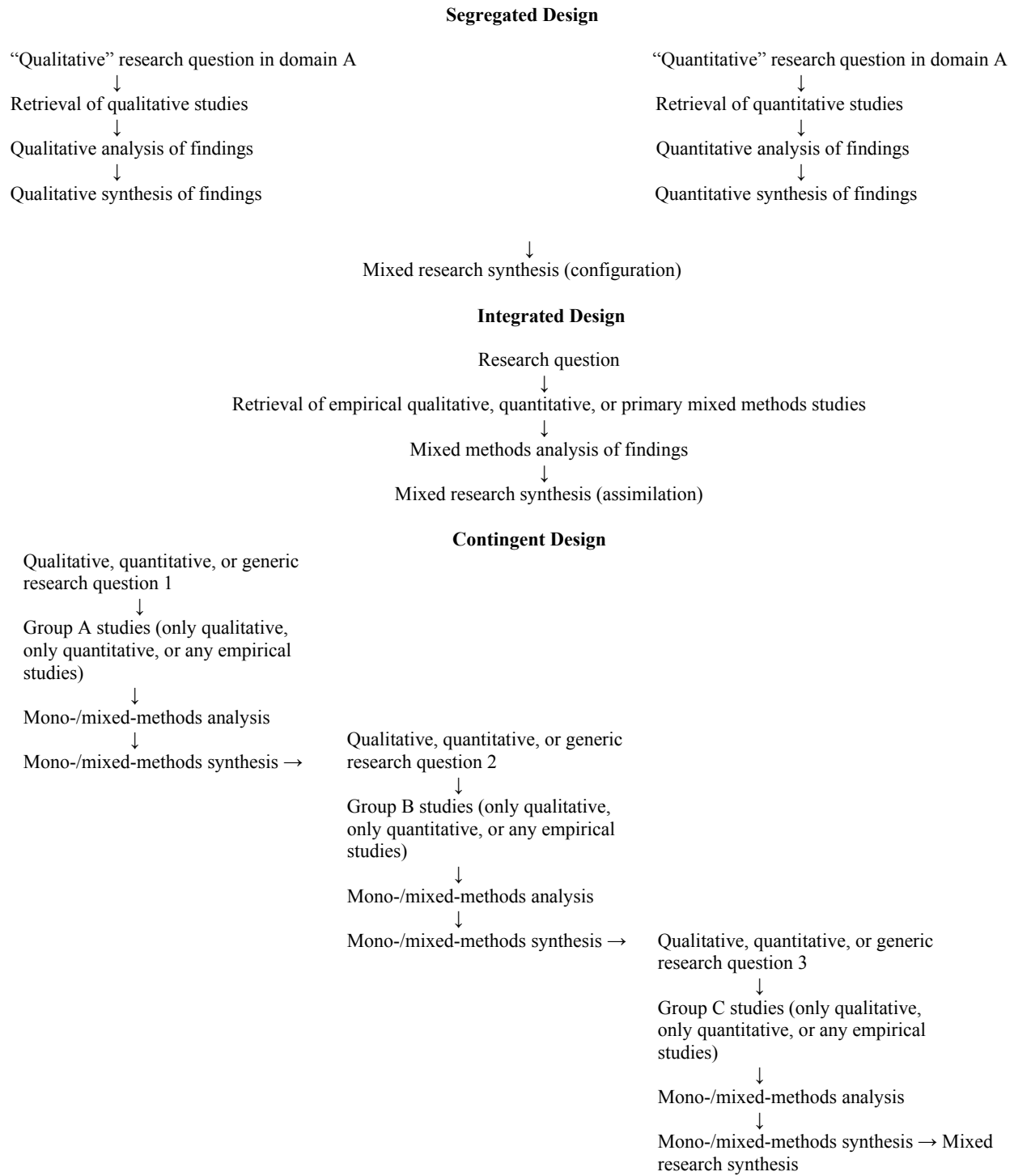
An example of complementarity is when a set of qualitative studies indicates that caring for children influences antiretroviral adherence in women and a set of quantitative studies indicates that sex influences adherence. The one finding neither confirms nor refutes the other as they are not referring to the same phenomenon and, therefore, are not subject to convergent validation. But the quantitatively-produced finding that sex predicted adherence may be clarified or explained by the qualitatively-produced finding that caring for children influenced adherence because childrearing is a gender-marked obligation, or a responsibility culturally prescribed for women as opposed to men. Similarly, the quantitatively-produced finding that race predicted non-adherence may be clarified or explained by the qualitatively-produced finding that HIV-positive African Americans’ knowledge of the Tuskegee syphilis study was a recurring reason given for their not taking antiretroviral drugs, as they saw these drugs as reprising the experimental and genocidal imperatives of the Tuskegee study.

These examples suggest that qualitative and quantitative research findings are complementary in linking causal explanations to causal observations (Maxwell, 2004a, 2004b). Quantitative findings indicate *that-knowledge* (e.g., that being female and being African American led to lower levels of adherence in comparison to being male or being white), while qualitative findings indicate *why-knowledge*, or the gender or race performances or relations that might explain these observations.

*Configuration vs. assimilation of findings.* The segregated design is also the design of choice when mixed research synthesis is conceived as the configuration, as opposed to assimilation, of qualitative and quantitative research findings. The configuration of findings is the arrangement of complementary findings into a line of argument (Noblit & Hare, 1988), a theory that posits relationships among concepts, or a narrative that posits a temporal ordering of events (e.g., Greenhalgh et al., 2005; Pound et al., 2005). Because qualitative and quantitative findings are viewed as addressing different aspects of a target phenomenon (i.e., as in a complementary relationship), they are also viewed as resistant to direct assimilation into each other. Unlike findings across studies seen to address the same relationship or aspect of a phenomenon, findings conceived as complementary cannot be reduced. Instead, they can only be organized into a coherent whole. For example, qualitative findings may be positioned as antecedent, mediating, or moderating



Table 1  
*Designs for Mixed Research Synthesis Studies*



variables explaining or delineating the conditions for the occurrence of events depicted in quantitative findings. Alternatively, quantitative findings can be used to make more explicit comparisons between groups only implied in qualitative findings.

#### *Integrated Design*

In the integrated design shown in Table 1, the methodological differences between qualitative and quantitative studies are minimized as both kinds of studies are viewed as producing findings that can readily be transformed into each other. This design is based on the assumptions that: (a) any differences between qualitative and quantitative studies that do exist do not warrant separate analyses and syntheses of their findings; (b) studies designated as qualitative or quantitative are not necessarily distinguishable from each other; (c) both qualitative and quantitative studies in a common research domain can address the same research purposes and questions; and (d) syntheses of both qualitative and quantitative findings can be produced from methods developed for qualitative and quantitative findings. The integrated design is most appropriate when: (a) qualitative and quantitative findings in a designated body of research are viewed as able to confirm, extend, or refute each other and when (b) mixed research synthesis is defined as the assimilation (as opposed to configuration) of research findings.

In integrated designs, the studies in a targeted domain are grouped for synthesis not by methods (i.e., qualitative and quantitative), but rather by findings viewed as answering the same research questions, or addressing the same aspects of a target phenomenon. Here findings addressing the same aspects may extend each other, which can be seen as a form of confirmation. An example of extension is when one set of findings indicates that having to take a large number of pills is a reason for not adhering to antiretroviral therapy, while another set of findings specifies the number of pills below which HIV-positive persons generally adhere and above which few persons adhere.

Mixed research synthesis is accomplished by mixed methods analysis. The analytic emphasis is on transforming findings to enable them to be combined. Transformation includes *qualitizing*, or converting quantitative findings into qualitative form so that they can be combined with other qualitative data and subjected to qualitative analysis, and *quantitizing*, or converting qualitative findings into quantitative form so that they can be combined with other quantitative data and subjected to quantitative analysis (Onwuegbuzie & Teddlie, 2003). Because few mixed research synthesis efforts exist, and because few reported qualitative or quantitative research syntheses have derived from the use of transformation

techniques, instances of *qualitizing* and *quantitizing* have appeared largely in reports of primary mixed methods research. Accordingly, examples of *quantitizing* with which we are experimenting in our on-going methods study that hold promise for mixed research synthesis include the calculation of effect sizes of qualitative findings and the translation of qualitatively-produced themes into predictor variables (Onwuegbuzie, 2003; Sandelowski & Barroso, in press). Examples of *qualitizing* that we are experimenting with include the conversion of quantitatively-produced correlations to themes, typologies, or case profiles (Onwuegbuzie & Teddlie, 2003; Tashakkori & Teddlie, 1998). Methods proposed to bridge the case-intensive world of qualitative research with the variable-extensive world of quantitative research include Ragin's (1987, 2000) qualitative comparative case method and Bayesian approaches to meta-analysis (Howard, Maxwell, & Fleming, 2000; Roberts, Dixon-Woods, Fitzpatrick, Abrams, Jones, 2002).

#### *Contingent Design*

In the contingent design shown in Table 1, the results of synthesizing the findings in a designated group of studies to answer one research question determine the next group of studies that will be retrieved and analyzed to answer a second research question the results of which, in turn, may lead to the analysis of a third group of studies retrieved to answer yet another research question. The cycle of systematic review continues until a comprehensive research synthesis can be presented that addresses researchers' objectives. For example, an initial focus on all reports of findings of studies testing the effectiveness of interventions to promote antiretroviral adherence in women might lead to subsequent searches for reports of findings concerning gender differences in antiretroviral adherence, the experience of HIV infection in women, or concerning HIV-related stigma to explain how and/or why these interventions succeeded or failed.

Contingent designs may or may not depend on hard lines drawn between qualitative and quantitative studies and between qualitative and quantitative methods of research synthesis. Contingent designs may be more like segregated designs in posing a series of research questions conceived to be amenable only to qualitative or quantitative studies, each set of which are analyzed with qualitative or quantitative methods, respectively, to produce the qualitative and quantitative research syntheses that will ultimately be configured into a theoretical or narrative rendering of findings. Alternatively, contingent designs may be more like integrated designs in posing a series of research questions deemed answerable by both qualitative and

quantitative studies in a targeted domain of research. The findings of these studies can then be assimilated. In short, the defining feature of contingent designs is the cycle of research synthesis studies conducted to answer questions raised by previous syntheses, not the grouping of studies or methods as qualitative and quantitative.

#### The Future of Mixed Research Synthesis

The viability of the mixed research synthesis enterprise rests on finding ways to make the seemingly incomparable comparable in order to make the seemingly uncombinable combinable (i.e., assimilable, arrangeable, or some other process of putting or using qualitative and quantitative findings together). Given the complexity of these goals, it is not surprising that mixed research synthesis methods have yet to be developed that satisfactorily accommodate the singularity, descriptive precision, and intricacy of qualitative research findings and the generality, numerical precision, and single-dimensionality of quantitative research findings (Buchanan, 1992; Sivesind, 1999). Although much work remains before mixed research synthesis can secure its place in the repertoires of mixed methods research and systematic review, the effort is well worth it as it has the potential to enhance both the significance and utility for practice of the many qualitative and quantitative research studies constituting shared domains of research.

The lead editors for this article were R. Burke Johnson and Anthony J. Onwuegbuzie.

#### References

- Barbour, R. S. (2000). The role of qualitative research in broadening the "evidence base" for clinical practice. *Journal of Evaluation in Clinical Practice*, 6, 155-163.
- Barbour, R. S., & Barbour, M. (2003). Evaluating and synthesizing qualitative research: The need to develop a distinctive approach. *Journal of Evaluation in Clinical Practice*, 9, 179-186.
- Booth, A. (2001, May). *Cochrane or cock-eyed? How should we conduct systematic reviews of qualitative research?* Paper presented at the Qualitative Evidence-Based Practice Conference: Taking a Critical Stance, Coventry University. Retrieved December 24, 2005, from *Education-Line* at <http://www.leeds.ac.uk/educol/documents/00001724.doc>
- Buchanan, D. R. (1992). An uneasy alliance: Combining qualitative and quantitative research methods. *Health Education Quarterly*, 19, 117-135.
- Campbell, R., Pound, P., Pope, C., Britten, N., Pill, R., Morgan, M., et al. (2003). Evaluating meta-ethnography: A synthesis of qualitative research on lay experiences of diabetes and diabetes care. *Social Science & Medicine*, 56, 671-684.
- Clarke, J. B. (1999). Evidence-based practice: A retrograde step? The importance of pluralism in evidence generation for the practice of health care. *Journal of Clinical Nursing*, 8, 89-94.
- Conn, V. S., & Rantz, M. J. (2003). Research methods: Managing primary study quality in meta-analyses. *Research in Nursing & Health*, 26, 322-333.
- Cooper, H. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 209-240). Thousand Oaks, CA: Sage.
- Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. London: Sage.
- Dixon-Woods, M., Agarwal, S., Young, B., Jones, D., & Sutton, A. (2004). *Integrative approaches to qualitative and quantitative evidence*. London: Health Development Agency. Retrieved November 13, 2005, from [http://www.med.umich.edu/csp/Course%20materials/Summer%202005/Goold\\_Integrative\\_Approaches.pdf](http://www.med.umich.edu/csp/Course%20materials/Summer%202005/Goold_Integrative_Approaches.pdf)
- Dixon-Woods, M., Fitzpatrick, R., & Roberts, K. (2001). Including qualitative research in systematic reviews: Opportunities and problems. *Journal of Evaluation in Clinical Practice*, 7, 125-133.
- Doyle, L. H. (2003). Synthesis through meta-ethnography: Paradoxes, enhancements, and possibilities. *Qualitative Research*, 3, 321-345.
- Eisner, E. W. (1991). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice*. New York: Macmillan.
- Estabrooks, C. A. (1999). Will evidence-based nursing practice make practice perfect? *Canadian Journal of Nursing Research*, 30, 273-294.

- Evans, D. (2003). Hierarchy of evidence: A framework for ranking evidence evaluating health care interventions. *Journal of Clinical Nursing, 12*, 77-84.
- Evans, D., & Pearson, A. (2001). Systematic reviews of qualitative research. *Clinical Effectiveness in Nursing, 5*, 111-119 (with commentaries and author response).
- Forbes, A., & Griffiths, P. (2002). Methodological strategies for the identification and synthesis of "evidence" to support decision-making in relation to complex healthcare systems and practices. *Nursing Inquiry, 9*, 141-155.
- Freese, L. (1980). The problem of cumulative knowledge. In L. Freese (Ed.), *Theoretical methods in sociology* (pp. 13-69). Pittsburgh, PA: University of Pittsburgh Press.
- Giacomini, M. K. (2001). The rocky road: Qualitative research as evidence. *Evidence-Based Medicine, 6*, 4-6.
- Glass, G. V. (2000). *Meta-analysis at 25*. Retrieved December 24, 2005, from <http://glass.ed.asu.edu/gene/papers/meta25.html>
- Glasziou, P. P., & Sanders, S. L. (2002). Investigating causes of heterogeneity in systematic reviews. *Statistics in Medicine, 21*, 1503-1511.
- Green, J., & Britten, N. (1998). Qualitative research and evidence based medicine. *British Medical Journal, 316*, 1230-1232.
- Greenhalgh, T. (2002). Integrating qualitative research into evidence based practice. *Endocrinology & Metabolism Clinics of North America, 31*, 583-601.
- Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., Kyriakidou, O., & Peacock, R. (2005). Storylines of research in diffusion of innovation: A meta-narrative approach to systematic review. *Social Science & Medicine, 61*, 417-430.
- Guba, E. G., & Lincoln, Y. S. (2005). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed., pp. 191-215). Thousand Oaks, CA: Sage.
- Gupta, M. (2003). A critical appraisal of evidence-based medicine: Some ethical considerations. *Journal of Evaluation in Clinical Practice, 9*, 111-121.
- Hammersley, M. (2001). On "systematic" reviews of research literatures: A "narrative" response to Evans & Benefield. *British Educational Research Journal, 27*, 543-554.
- Hampton, J. R. (2002). Evidence-based medicine, opinion-based medicine, and real-world medicine. *Perspectives in Biology and Medicine, 45*, 549-568.
- Harden, A., & Thomas, J. (2005). Methodological issues in combining diverse study types in systematic reviews. *Internal Journal of Social Research Methodology, 8*, 257-271.
- Hawker, S., Payne, S., Kerr, C., Hardey, M., & Powell, J. (2002). Appraising the evidence: Reviewing the disparate data systematically. *Qualitative Health Research, 12*, 1284-1299.
- Higgins, J. P. T., & Green, S. (Eds.), (2005, May). *Cochrane handbook for systematic reviews of interventions 4.2.5*. Retrieved December 24, 2005, from <http://www.cochrane.org/resources/handbook/hbook.htm>
- Howard, G. S., Maxwell, S. E., & Fleming, K. J. (2000). The proof of the pudding: An illustration of the relative strengths of null hypothesis, meta-analysis, and Bayesian analysis. *Psychological Methods, 5*, 315-332.
- Hunt, M. (1997). *How science takes stock*. New York: Russell Sage Foundation.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*, 14-26.
- Jones, K. (2004). Mission drift in qualitative research, or moving toward a systematic review of qualitative studies, moving back to a more systematic narrative review. *The Qualitative Report, 9*(1), 95-112. On line journal, retrieved August 29, 2004, from [www.nova.edu/ssss/QR/QR9-1/jones.pdf](http://www.nova.edu/ssss/QR/QR9-1/jones.pdf)
- Lemmer, B. Grellier, R., & Steven, J. (1999). Systematic review of nonrandom and qualitative research literature: Exploring and uncovering an evidence base for health visiting and decision making. *Qualitative Health Research, 9*, 315-328.
- Leys, M. (2003). Health care policy: Qualitative evidence and health technology assessment. *Health Policy, 65*, 217-226.
- MacLure, M. (2005). "Clarity bordering on stupidity": Where's the quality in systematic review? *Journal of Education Policy, 20*, 393-416.
- Madjar, I., & Walton, J. A. (2001). What is problematic about evidence? In J. M. Morse, J. M. Swanson, & A. J. Kuzel (Eds.), *The nature of qualitative evidence* (pp. 28-45). Thousand Oaks, CA: Sage.
- Matt, G. E. (2003). Will it work in Munster? Meta-analysis and the empirical generalization of causal relationships. In R. Schulze, H. Holling, & D. Bohning (Eds.), *Meta-analysis:*

- New developments and applications in medical and social sciences* (pp. 113-139). Cambridge, MA: Hogrefe & Huber.
- Maxwell, J. A. (2004a). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher*, 33, 3-11.
- Maxwell, J. A. (2004b). Using qualitative methods for causal explanation. *Field Methods*, 16, 243-264.
- Mays, N., Pope, C., & Popay, J. (2005). Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. *Journal of Health Services Research and Policy*, 10(Suppl. 1), 6-20.
- McKenna, H., Cutcliffe, J., & McKenna, P. (1999). Evidence-based practice: Demolishing some myths. *Nursing Standard*, 14, 39-42.
- Miller, S., & Fredericks, M. (2003). The nature of "evidence" in qualitative research methods. *International Journal of Qualitative Methods*, 2(1). Article 4. Retrieved September 22, 2003, from [http://www.ualberta.ca/~iiqm/backissues/2\\_1/html/miller.html](http://www.ualberta.ca/~iiqm/backissues/2_1/html/miller.html)
- Mitchell, G. J. (1999). Evidence-based practice: Critique and alternative view. *Nursing Science Quarterly*, 12, 30-35.
- Mulrow, C., Langhorne, P., & Grimshaw, J. (1997). Integrating heterogeneous pieces of evidence in systematic reviews. *Annals of Internal Medicine*, 127, 989-995.
- Mykhalovskiy, E., & Weir, L. (2004). The problem of evidence-based medicine: Directions for social science. *Social Science & Medicine*, 59, 1059-1069.
- Noblit, G. W., & Hare, R. D. (1988). *Meta-ethnography: Synthesizing qualitative studies*. Newbury Park, CA: Sage.
- Onwuegbuzie, A. J. (2003). Effect sizes in qualitative research: A prolegomenon. *Quality & Quantity*, 37, 393-409.
- 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 December 1, 2005, from <http://www.nova.edu/ssss/QR/QR9-4/onwuegbuzie.pdf>.
- 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.
- Paterson, B. L., Thorne, S. E., Canam, C., & Jillings, C. (2001). *Meta-study of qualitative health research: A practical guide to meta-analysis and meta-synthesis*. Thousand Oaks, CA: Sage.
- Pellegrino, E. D. (2002). Medical evidence and virtue ethics: A commentary on Zarkovich and Upshur. *Theoretical Medicine*, 23, 397-402.
- Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: Horses for courses. *Journal of Epidemiology & Community Health*, 57, 527-533.
- Popay, J., & Roen, K. (2003, May 23). *Synthesis of evidence from research using diverse study designs: A preliminary review of methodological work*. Unpublished report for the Social Care Institute of Excellence, Lancaster University, UK.
- Popay, J., & Williams, G. (1998). Qualitative research and evidence-based health care. *Journal of the Royal Society of Medicine*, 91(Suppl. 35), 32-37.
- Pope, C., & Mays, N. (1995). Reaching the parts other methods cannot reach: An introduction to qualitative methods in health and health services research. *British Medical Journal*, 311, 42-45.
- Pound, P., Britten, N., Morgan, M., Yardley, L., Pope, C., Daker-White, G., et al. (2005). Resisting medicines: A synthesis of qualitative studies of medicine taking. *Social Science & Medicine*, 6, 133-155.
- Ragin, C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.
- Ragin, C. (2000). *Fuzzy-set social science*. Chicago: University of Chicago Press.
- Ray, L. D., & Mayan, M. (2001). Who decides what counts as evidence? In J. M. Morse, J. M. Swanson, & A. J. Kuzel (Eds.), *The nature of qualitative evidence* (pp. 50-73). Thousand Oaks, CA: Sage.
- Roberts, K. A., Dixon-Woods, M., Fitzpatrick, R., Abrams, K. R., & Jones, D. R. (2002). Factors affecting uptake of childhood immunization: A Bayesian synthesis of qualitative and quantitative evidence. *The Lancet*, 360, 1596-1599.
- Rolfe, G. (2002). Faking a difference: Evidence-based nursing and the illusion of diversity. *Nurse Education Today*, 22, 3-12.
- Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal*, 312(7023), 71-72.

- Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-based medicine: How to practice and teach EBM*. Edinburgh: Churchill Livingstone.
- Sandelowski, M. (2004). Using qualitative research. *Qualitative Health Research, 14*, 1366-1386.
- Sandelowski, M., & Barroso, J. (2003). Motherhood in the context of maternal HIV infection. *Research in Nursing & Health, 26*, 470-482.
- Sandelowski, M., & Barroso, J. (2005). The travesty of choosing after positive prenatal diagnosis. *JOGNN: Journal of Obstetric, Gynecologic, and Neonatal Nursing, 34*, 307-318.
- Sandelowski, M., & Barroso, J. (in press). *Handbook for synthesizing qualitative research*. New York: Springer.
- Sandelowski, M., Docherty, S., & Emden, C. (1997). Qualitative metasynthesis: Issues and techniques. *Research in Nursing & Health, 20*, 365-371.
- Sandelowski, M., Voils, C. I., & Barroso, J. (2006). *Comparability work and the management of difference in mixed research synthesis*. Manuscript submitted for publication.
- Schulze, R., Holling, H., & Bohning, D. (Eds.). (2003). *Meta-analysis: New developments and applications in medical and social sciences*. Cambridge, MA: Hogrefe & Huber.
- Sivesind, K. H. (1999). Structured, qualitative comparison: Between singularity and single-dimensionality. *Quality & Quantity, 33*, 361-380.
- Slavin, R. E. (1995). Best evidence synthesis: An intelligent alternative to meta-analysis. *Journal of Clinical Epidemiology, 48*, 9-18.
- Stangl, D. K., & Berry, D. A. (Eds.). (2000). *Meta-analysis in medicine and health policy*. New York: Marcel Dekker.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: 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.
- Thorne, S., Jensen, L., Kearney, M. H., Noblit, G., & Sandelowski, M. (2004). Qualitative meta-synthesis: Reflections on methodological orientation and ideological agenda. *Qualitative Health Research, 14*, 1342-1365.
- Timmermans, S., & Berg, M. (2003). *The gold standard: The challenge of evidence-based medicine and standardization in health care*. Philadelphia, PA: Temple University Press.
- Traynor, M. (2002). The oil crisis, risk and evidence-based practice. *Nursing Inquiry, 9*, 162-169.
- Trinder, L. (2000). A critical appraisal of evidence-based practice. In L. Trinder & S. Reynolds (Eds.), *Evidence-based practice: A critical appraisal* (pp. 212-241). Oxford, UK: Blackwell Science.
- Trinder, L., & Reynolds, S. (Eds.). (2000). *Evidence-based practice: A critical appraisal*. Oxford, UK: Blackwell Science.
- Upshur, R. E. (2001). The status of qualitative research as evidence. In J. M. Morse, J. M. Swanson, & A. J. Kuzel (Eds.), *The nature of qualitative evidence* (pp. 5-26). Thousand Oaks, CA: Sage.
- Weinholtz, D., Kacer, B., & Rocklin, T. (1995). Salvaging quantitative research with qualitative data. *Qualitative Health Research, 5*, 388-397.

Notes

<sup>1</sup>This study, entitled *Integrating qualitative & quantitative research findings*, is funded by the National Institute of Nursing Research, National Institutes of Health, 5 R01 NR004907, June 3, 2005-March 31, 2010. Support was also provided by Career Development Award MRP 04-216-1 awarded to Corrine I. Voils from the Health Services Research and Development Service of the Department of Veterans Affairs. The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

<sup>2</sup>Both systematic review and evidence-based practice have generated strong and even scathing criticism. See, for example, Clarke, 1999; Estabrooks, 1999; Gupta, 2003; Hampton, 2002; MacLure, 2005; Mykhalovskiy & Weir, 2004; Pellegrino, 2002; Rolfe, 2002; Timmermans & Berg, 2003; Traynor, 2002; & Trinder, 2000.

<sup>3</sup>Differences even exist in categorizing these differences; for example, what we refer to here as paradigms others refer to as theories.

<sup>4</sup>Having ascertained what appears to be a refutation, or a contradictory view of the same relationship, researchers will likely want to ascertain conditions related to both the target phenomenon and to the nature of the research itself that might explain why

depression influenced adherence in one group of studies and had no influence in a second group of studies. That is, they may want to ascertain whether an apparent refutation is an actual refutation. Similarly, apparent confirmations may not be actual confirmations when reviewers probe findings that seem to replicate each other. Confirmation and refutation are processes not so simple as we depict them here and as they are typically presented in the research synthesis literature. We plan to address the complexities of these processes in a future paper.

## Mixed Methods Research: Are the Methods Genuinely Integrated or Merely Parallel?<sup>1</sup>

Robert K. Yin  
COSMOS Corporation

*Using mixed methods within the confines of a single study can simultaneously broaden and strengthen the study. However, a continuing challenge is to maintain the integrity of the single study, compared to inadvertently permitting the study to decompose into two or more parallel studies. This article therefore deliberately focuses on those issues central to integrating mixed methods within a single study. The article points to pitfalls and remedies for integrating mixed methods throughout the conduct of a single study, ranging from the ways of splitting the study's initial research questions to the strategies for conducting analyses. The stronger the "mix" of methods throughout these procedures the more that researchers can derive the benefits from using mixed approaches.*

This article briefly discusses specific pitfalls and reminders in designing and conducting mixed methods research. Ample arrays of mixed methods procedures exist elsewhere (e.g., Tashakkori & Teddlie, 1998; and the other articles in this volume). However, this article deliberately focuses on those issues central to integrating mixed methods within a single study. Without such integration, different methods may sit in parallel, potentially leading to multiple studies, and not the desired "mixing" of methods implicit in mixed methods research.

### Using Mixed Methods Within the Confines of a "Single" Study

Mixed methods research plays an important if not essential role in educational research. Some investigators, including the guest editor of the present volume, have proposed mixed methods research as a "research paradigm whose time has come" (Johnson & Onwuegbuzie, 2004, p. 14). New paradigm or not, mixed methods research has been conducted in education research for a long time (e.g., Sieber, 1973).

Johnson and Onwuegbuzie (2004) provide the point of departure for the present article. They define mixed methods research as "the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts, or language into a *single* study [*italics added*]" (p. 17). The focus on a *single* study is critical to mixed methods research. Implicit in the prominent role played by a single study is the valuing of mixed methods in producing

converging evidence, presumably more compelling than might have been produced by any single method alone.

In contrast, if a research effort consists of multiple, related studies rather than a single study, little distinctive contribution arises from attending to the use of multiple methods. In fact, when investigators have permitted a research effort to decompose into multiple studies, mixed methods research may not have taken place at all.

Thus, forcing the use of multiple methods into the confines of a single study means avoiding a more traditional realm whereby separate studies have been conducted and later synthesized. Under that circumstance, a common cross-study question might be whether the findings from the separate studies had been confirmatory, replicative, or contradictory. But the cross-study analysis would resemble well-trodden paths involving research syntheses (e.g., Cooper & Hedges, 1994), meta-analyses (e.g., Lipsey & Wilson, 2001), or similar aggregative procedures. They all fall outside of (and would not be called) mixed methods research.<sup>2</sup>

### Mixed Methods: Relevant Combinations Go Beyond the Quantitative-Qualitative Dichotomy

At the same time, this article goes beyond the initial definition by Johnson and Onwuegbuzie (2004) as well as other related works (e.g., Tashakkori & Teddlie, 1998) in one respect: Mixed methods research embraces much more than the traditional dichotomy between qualitative and quantitative research. Such a dichotomy once threatened to split the evaluation community (e.g., Lincoln, 1991; Reichardt & Rallis, 1994; Sechrest, 1992; Yin, 1994).

The dichotomous view masks the reality that there can be many different "mixes" or combinations of methods. For instance, in education research, some combination of experimentation and surveys—both being

---

Correspondence for this article should be addressed to Robert K. Yin, COSMOS Corporation, 3 Bethesda Metro Center, Suite 400, Bethesda, MD 20814. E-mail: ryin@cosmoscorp.com

---



forms of “quantitative” methods—might deal better with the dual needs of addressing internal and external validity than either method alone (e.g., Berends & Garet, 2002). Thus, a single study combining experimentation and surveys would be an example of mixed methods research even though no qualitative method had been included in the study.

Likewise, scholars have long recognized different forms of “qualitative” methods (e.g., Guba & Lincoln, 1994; Trow, 1957), which also could be combined into a single, mixed methods study even though no quantitative method had been included. Finally, even within experimental methods the mixing of “group” designs and research (e.g., Campbell & Stanley, 1966) and “single-subject” research (e.g., Hersen & Barlow, 1976) could produce a powerful, single study of human behavior.

Once freed from the quantitative-qualitative dichotomy, the relevance and reality of a broad variety of “mixes” emerges. The broad variety recognizes the true diversity of the research methods used in education, ranging from clinical research to econometrics to experiments. The broad variety also recognizes the variants within specific methods, such as in-person and mail surveys, ethnographic and case study fieldwork, and laboratory and “natural” experiments. Figure 1 enumerates the broad diversity of methods available in education research. Referencing this exhibit, mixed methods research can exist whenever a single study includes methods *between* or—but to a much lesser extent—*within* the five circled items.

A review of these combinations and their potential utility, citing exemplary examples from published research, might make an important contribution to education research. However, the present article has a more modest goal. The article focuses mainly on one question: How to tighten the use of mixed methods so that they do in fact occur as part of a *single* study. The article briefly discusses the needed integration in carrying out five procedures.<sup>3</sup> These include the relationship among the mixed methods with regard to a study’s:

1. Research questions
2. Units of analysis
3. Samples for study
4. Instrumentation and data collection methods
5. Analytic strategies

The claim is that, the more that a single study integrates mixed methods across these five procedures, the more that mixed methods research, as opposed to multiple studies, is taking place.

#### Research Questions

A single study no doubt starts with a single set of research questions. A frequent practice in allegedly carrying out mixed methods research is to split the original set of questions, so that different research methods address different questions.

In education evaluations, a typical split is for qualitative methods to address “process” questions and

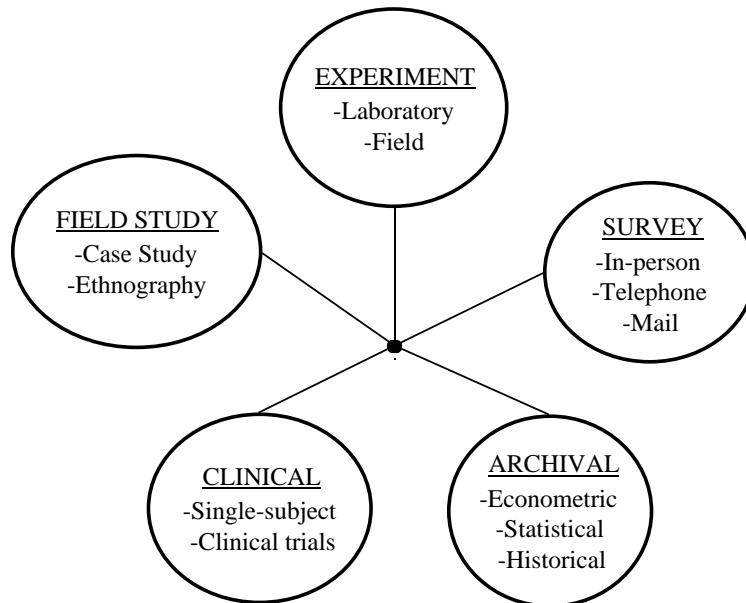


Figure 1. Mixed Methods Research: Broad Variety of “Mixes,” Beyond Simple Qualitative-Quantitative Combinations

MIXED METHODS RESEARCH: ARE THE METHODS GENUINELY INTEGRATED OR MERELY PARALLEL?

for quantitative methods to address “outcome” questions. Though this does not necessarily lead to the conduct of multiple studies, observe how a study would be strengthened if both the quantitative and qualitative methods each addressed some aspect of both process and outcome questions. Figure 2 illustrates both splits. A “process”-“outcome” split between, say, the survey and case study methods in an education evaluation potentially leads to separate studies (or “sub-studies”). The more desirable split has both the survey and case study efforts each addressing both “process” and “outcome” questions.

How research questions may be split or decomposed requires a lengthy discussion beyond the scope of this article. The simple suggestion is that the sharper the splits among the different research methods, the more readily a research effort, initially construed as a single study, might become a set of multiply-related studies.

Units of Analysis

Every study has an implicit if not explicit unit of analysis or assignment. The unit of analysis, though not necessarily the unit of data collection, holds a study together. The challenge in using mixed methods arises because different methods inherently favor different units of analysis—leading to another threat to the integrity of a single study.

Figure 3 again shows two contrasting conditions. The right side of this figure depicts a possibly classic example of an appropriately integrated study: a case study of a labor union, published 50 years ago by three eminent social scientists (Lipset, Trow, & Coleman, 1956). The study collected data at every level of this complex organization, using a variety of (quantitative

and qualitative) methods. Despite the varied methods, the researchers consistently maintained the same point of reference—which was the labor union as a single organization. Such persistent reference to this unit of analysis created the needed integrative force to blend all of the methods into a single study, eventually reported in the form of a 450-page book.

By comparison, the left side of Figure 3 depicts a prominent research investigation into community health care, illustrating how a single research effort decomposed into two isolated studies. The investigators’ original goal was to gain insight into the providers and clients of the set of health maintenance organizations (HMOs) in a particular community (Ginsburg, 1996). The multiple methods called for field studies that collected data from the set of HMOs, and a random-digit-dial (RDD) survey that covered a sample of community residents. However, the survey sample, not surprisingly, captured few clients, and the fieldwork, also not surprisingly, favored information about the service providers but not the clients. In fact, different units of analyses were inadvertently at work: by dint of its RDD design, the survey had defined a geographic area as the unit of analysis, whereas the fieldwork had defined a service delivery system. As a result, the findings from the two methods could not be integrated, leaving the original desire to study provider-client relationships unfulfilled.

Education research risks such decomposing, too. A typical study might have fieldwork about school districts and survey data from teachers. To prevent the study from decomposing into two isolated studies, the fieldwork might include focus groups of teachers, deliberately covering some of the same questions as those in the survey (but integrating the focus group findings with the rest of the fieldwork). Similarly, the

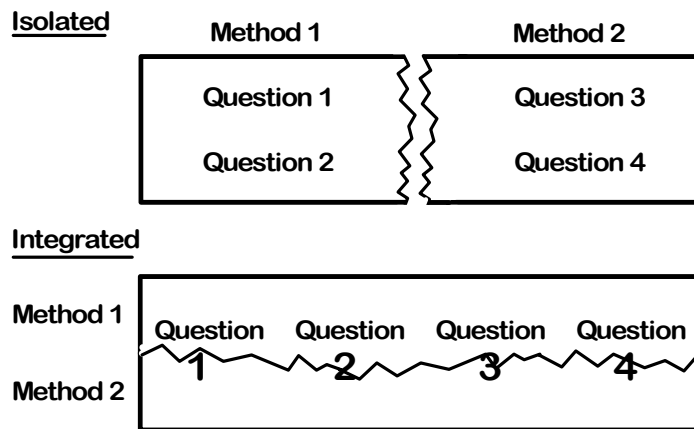


Figure 2. Research Questions: Isolated or Integrated?

survey sample might be stratified to reflect the district-related issues of study (e.g., different grade levels, subjects, or specialties, such as counseling), not just a representative sample of all teachers. In other words, each method can reach into the realm of the other, to produce a single mixed methods study.

Samples

Sampling procedures also need to be considered carefully in maintaining a single study while using mixed methods. Most desirably, the samples of each method may be nested within that of the other.

The nesting may be in either direction. For instance, fieldwork samples may be nested within survey samples, as in the by now routine situation where case studies are conducted on a small set of schools that are part of a much larger survey sample of schools. At COSMOS, such a sampling arrangement is part of an ongoing study of comprehensive school reform.

At the same time, the nesting can be in the other direction, too. For instance, in another study, COSMOS is evaluating public school choice initiatives. The fieldwork focuses on 13 districts that received special federal funding. Within each of these districts, a school survey deliberately covers those schools believed by the district to have participated in its choice initiative.

Figure 4 shows these two nesting patterns. They are but two of many combinations that might be pertinent for maintaining a single study while using mixed methods.

Instrumentation and Data Collection

Different methods tend to use different types of instruments. As examples, experiments typically use some kind of apparatus, including devices for recording behavior; surveys use questionnaires; case studies use field protocols; and ethnographies might use participant-observation.

Mixed methods research cannot change this array, as each method’s preferred instrumentation is central to the method itself. However, despite the differences, the various instruments could contain directly analogous variables, if not actual items. The more that the items overlap or complement each other; the more that the mixed methods can be part of a single study. Conversely, greater divergence can again lead to multiple studies.

Within each of the two previously mentioned education evaluations underway at COSMOS—one examining public school choice and the other comprehensive school reform—the items in the various instruments were cross-walked. Some of the items were numeric (e.g., using district fieldwork documentation to define the number of students participating in a choice initiative, and having the same item asked in the school surveys). Other items were more conceptual and qualitative (e.g., using fieldwork to define the “comprehensiveness” of a whole-school reform initiative, and using an array of survey items to cover the same “comprehensiveness” concepts).

Figure 3. Unit of Analysis: Isolated or Integrated?

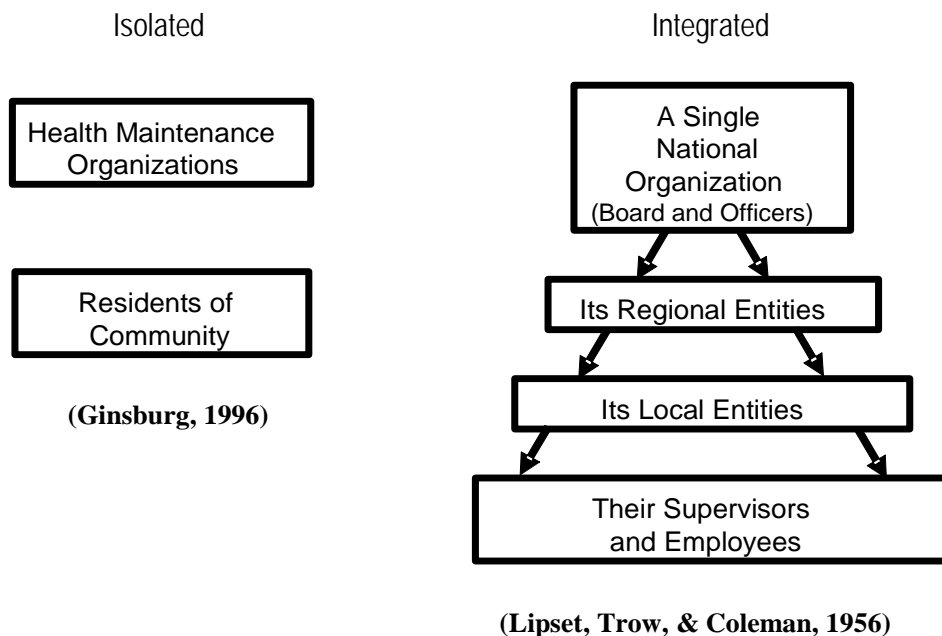
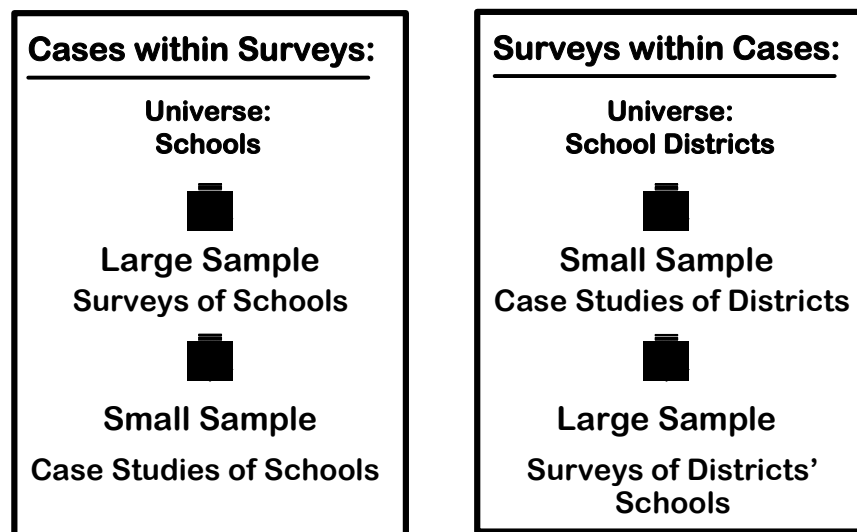


Figure 4: Samples: Integration Through Nested Arrangements



Although the measures were not exactly the same within each evaluation, the investigators deliberately tried to create directly comparable items, to assure the desired common scopes of data collection and variables. At the same time, the instruments were not limited to their common ground. By design, they also had many non-overlapping items. Figure 5 illustrates the desired cross-walking relationship when using different methods within the confines of a single study.

#### Analytic Strategies

Of all the procedures, analytic integration may be the trickiest of all. The mixed methods, if truly different methodologically, are likely to come with their own preferred and distinct analytic techniques. Under this circumstance, the goal is not to force the mixed methods into the exact same analytic routines. Rather, the goal is to design and carry out what might be called “counterpart” analyses. Such analyses should be formulated in directly analogous fashion, although they may use entirely different methodological techniques.

For instance, in one kind of formulation, mixed methods that are truly part of the same study can examine the relationships between the same dependent variable and associated independent variables. Such counterparts exist in the ongoing evaluation of comprehensive school reform, where the specifications of the survey’s regression models are mimicked by the analysis of “logic models” based on using case study methods. The analyses are counterparts in that both follow the same formulation—to study the student

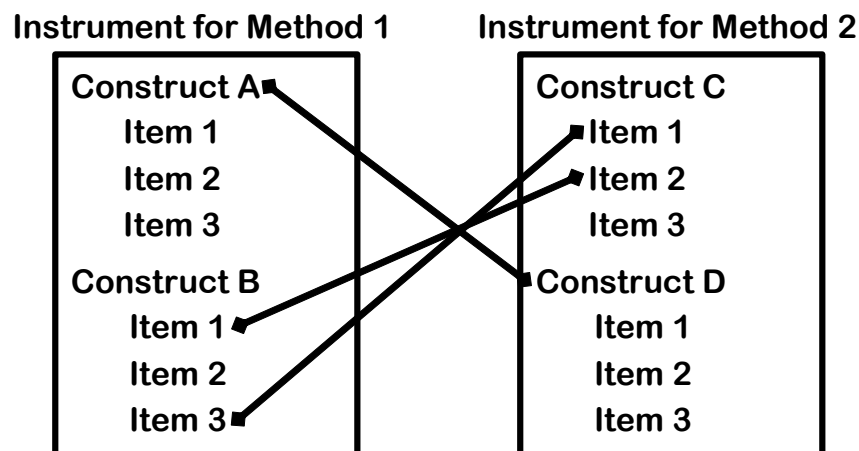
achievement trends that might be associated with schools that have been successfully reforming.

Another type of formulation disregards any relationship between dependent and independent variables and is more descriptive: Do the mixed methods at least tell the same story in describing either the dependent or independent variables alone? Such corroboration has been part of our evaluation of public school choice, where districts’ claims about schools’ choice participation from fieldwork data are juxtaposed with schools’ rendition of their own participation as expressed in a large-sample school survey.

As another variant of this descriptive formulation, do both the qualitative and quantitative (or the experimental and non-experimental) work suggests similar typologies of the subject being studied? Such counterparts have been part of our evaluation of comprehensive school reform, with the two main methods each trying to corroborate the same five-fold typology: Schools that are “reforming,” “on the way to reforming,” “just starting to reform,” “not reforming,” and “reforming, but with difficulty.”

These illustrations provide a start for thinking about yet other types of counterparts. The point is, if a counterpart relationship is entirely missing—as in the situation where two or more methods address wholly different dependent, independent, or descriptive variables—the mixed methods are likely to form separate studies, not a single study.

Figure 5. Instrumentation and Data Collection: Crosswalked Items and Variables



### Summary

Summarizing this article is simple. The design and conduct of a single study involves an array of readily understood procedures, regarding: the research questions being addressed, the definition of the units of analyses, the structure of the samples being studied, the instrumentation and data collected, and the analytic strategies. The more that two (or more!) methods have been integrated into each of these procedures, the stronger the “mix” of methods. Conversely, if each method uses its own isolated procedures, the result will be separate studies using different methods. Though the studies may be complementary, they will not really represent mixed methods research.

The lead editors for this article were R. Burke Johnson and Anthony J. Onwuegbuzie.

### References

- Berends, M., & Garet M. S. (2002). In (re)search of evidence-based school practices: Possibilities for integrating nationally representative surveys and randomized field trials to inform educational policy, *Peabody Journal of Education*, 77(4), 28-58.
- Campbell, D. T., & Stanley, J. (1966). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.
- Cooper, H., & Hedges, L. V. (Eds.). (1994). *The handbook of research synthesis*. New York, NY: Russell Sage Foundation.
- Ginsburg, P. B. (1996). The RWJF community snapshots study: Introduction and overview. *Health Affairs*, 15(2), 7-20.
- Guba, E., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Lincoln, Y. S. (Eds.), *Handbook of qualitative research* (pp. 105-117). Thousand Oaks, CA: Sage
- Hersen, M., & Barlow, D. H. (1976). *Single case experimental designs: Strategies for studying behavior change*. New York, NY: Pergamon.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Lincoln, Y. S. (1991). The arts and sciences of program evaluation, *Evaluation Practice*, 12(1), 1-7.
- Lipset, S.M., Trow, M., & Coleman, J. (1956). *Union democracy: The inside politics of the international typographical union*. New York, NY: Free Press.
- Lipsey, M., & Wilson D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Reichardt, C. S., & Rallis, S. F. (1994). Qualitative and quantitative inquiries are not incompatible: A call for a new partnership. In C. S. Reichardt & S. F. Rallis (Eds.), *The qualitative-quantitative debate: New perspectives* (pp. 85-92). San Francisco: Jossey-Bass.
- Sechrest, L. (1992). Roots: Back to our first generation. *Evaluation Practice*, 13(1), 1-7.
- Sieber, S. D. (1973). The integration of fieldwork and survey methods. *American Journal of Sociology*, 78(6), 1335-1359.

MIXED METHODS RESEARCH: ARE THE METHODS GENUINELY INTEGRATED OR MERELY  
PARALLEL?

- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.
- Trow, M. (1957). Comment on participant observation and interviewing: A comparison. *Human Organization, 16*, 33-35.
- Yin, R. K. (1994). Evaluation: A singular craft. In C. S. Reichardt & S. F. Rallis (Eds.), *The qualitative-quantitative debate: New perspectives* (pp. 71-84). San Francisco: Jossey-Bass.

Notes

<sup>1</sup> This article is based on a paper presented by the author at the 2006 Annual Meeting of the American Educational Research Association.

<sup>2</sup> The discussion here refers to the more traditional syntheses, which typically aggregate across individual studies that have used the same general method. However, the role of research syntheses can be enhanced to include the synthesis of studies using different methods (see Sandelowski, Voils, & Barroso in this volume).

<sup>3</sup> For the purpose of clarity, each procedure is discussed separately. In real-life study designs, the procedures may be sufficiently related that two or more of them work in tandem.

## The Validity Issue in Mixed Research

**Anthony J. Onwuegbuzie**  
*University of South Florida*

**R. Burke Johnson**  
*University of South Alabama*

*In quantitative research, the importance of validity has been long accepted. In qualitative research, discussions of validity have been more contentious and different typologies and terms have been produced. In mixed methods research, wherein quantitative and qualitative approaches are combined, discussions about “validity” issues are in their infancy. We argue that because mixed research involves combining complementary strengths and nonoverlapping weaknesses of quantitative and qualitative research, assessing the validity of findings is particularly complex; we call this the problem of integration. We recommend that validity in mixed research be termed legitimation in order to use a bilingual nomenclature. Tashakkori and Teddlie’s (2003, 2006) evaluation criteria frameworks involving the concept of inference quality are summarized. Although providing a framework for assessing legitimation in mixed research always will be incomplete, it is important to address several legitimation types that come to the fore as a result of combining inferences from the quantitative and qualitative components of the study into the formation of meta-inferences. Nine types of legitimation are described here in order to continue this emerging and important dialogue among researchers and methodologists.*

This paper is focused on validity in mixed methods research or what we refer to more broadly as *mixed research*. However, to understand the validity issue (i.e., quality) in mixed research, a brief review of some related discussions in quantitative and qualitative research will be helpful for orientation. Because these issues have been discussed elsewhere in great detail, we provide only brief summaries of those literatures, but first we want to make a few introductory comments about our general approach to research validity or quality.

We try to take a “middle of the road” position, seeing some truth and insight to be gained from multiple perspectives. Our approach is only one among many, and we recommend that readers examine additional perspectives as more work is carried out in this emerging area in mixed methods research as well as in the more traditional areas of qualitative and

quantitative research quality. The “validity” issue, at least as we use the term, is not about singular truths, and it certainly is not limited to quantitative measurement; rather, by validity we mean that a research study, its parts, the conclusions drawn, and the applications based on it can be of high or low quality, or somewhere in between. Research needs to be defensible to the research and practice communities for whom research is produced and used. The arbiters of research quality will be the research stakeholders, which means that the quality or validity issue can have subjective, intersubjective, and objective components and influences. At the same time, research is something about which we can “rationally” speak, and usually, after considering our external and our internal or epistemic standards, we can meaningfully assert that some research is of higher quality for certain purposes than is other research (Longino, 1990). Anthropology, sociology, and psychology teach us that communities, cultures, and various kinds of groupings (including communities of researchers) have some *shared* norms, practices, values, and beliefs.

We aim our sense of justification at the research community that sees many advantages to *sometimes* using both qualitative and quantitative research in their single or highly related sets of research studies. One of the exciting results of much mixed research is that in a

---

Correspondence should be addressed to Anthony J. Onwuegbuzie, Dept. of Educational Measurement and Research, College of Education, University of South Florida, 4202 East Fowler Ave., EDU 162, Tampa, Florida 33620.  
Email: [tonyonwuegbuzie@aol.com](mailto:tonyonwuegbuzie@aol.com)

---

single study practical questions can be addressed, different perspectives can be examined, and if well documented, practitioners can obtain some sense of what might be useful in their local situations. We do not want to *oversell* mixed research, however; the evidence will be in the results. If mixed research produces useful results over time, as well as useful theory, then progress will have been made. We agree with Kurt Lewin's statement that "There is nothing so practical as a good theory" (Lewin, 1952, p. 169), and we hope that all researchers, including mixed researchers, will attempt to produce *good* theories and other research works.

#### Validity in Quantitative Research

In quantitative research, discussions of "validity" have been common and the importance of validity has been long accepted, and this is well documented in the literature. Building on the seminal works of Campbell and Stanley (Campbell, 1957; Campbell & Stanley, 1963), and many others, Onwuegbuzie (2003) presented 50 different threats to internal and external validity that might occur at the research design/data collection, data analysis, and/or data interpretation stages of the quantitative research process. These threats are presented in Figure 1, in what was later called the *Quantitative Legitimation Model*. As illustrated in Figure 1, Onwuegbuzie identified 22 threats to internal validity and 12 threats to external validity at the research design/data collection stage of the quantitative research process. At the data analysis stage, 21 and 5 threats to internal validity and external validity were conceptualized, respectively. Finally, at the data interpretation stage, 7 and 3 threats to internal validity and external validity were identified, respectively. In Figure 2, Onwuegbuzie, Daniel, and Collins' (in press) have presented a schematic representation of instrument score validity, which also is provided here for review by interested readers.

Another very important work in validity in quantitative research is found in Shadish, Cook, and Campbell (2001). These authors continue to build on Campbell's earlier work and classify research validity into four major types: statistical conclusion validity, internal validity, construct validity, and external validity. Other selected seminal works showing the historical development of validity in quantitative research are summarized in the following references: American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999), Bracht and Glass (1968), Campbell (1957), Campbell and Stanley (1963), Cook and Campbell (1979), Messick (1989, 1995), and Smith and Glass (1987).

#### Validity in Qualitative Research

In the qualitative research paradigm, a primary focus is for researchers to capture authentically the lived experiences of people. As noted by Denzin and Lincoln (2005), "Such experience, it is argued, is created in the social text written by the researcher. This is the representational problem. It confronts the inescapable problem of representation, but does so within a framework that makes the direct link between experience and text problematic" (p. 19).<sup>2</sup> Denzin and Lincoln (2005) also 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. This problem asks, 'How are qualitative studies to be evaluated in the contemporary, poststructural moment?'" (pp. 19-20).<sup>3,4</sup> Part of their solution to the "validity issue" has been to reconceptualize traditional quantitative validity concepts and to use labels that are more acceptable to qualitative researchers (Lincoln & Guba, 1985, 1990). One set of criteria (Lincoln & Guba, 1985) includes the following types: 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).

Another useful classification for validity in qualitative research was provided by Maxwell (1992), who identified the following five types of validity in qualitative research: descriptive validity (i.e., factual accuracy of the account as documented by the researcher), interpretive validity (i.e., the extent to which an interpretation of the account represents an understanding of the perspective of the underlying group and the meanings attached to the members' words and actions), theoretical validity (i.e., the degree to which a theoretical explanation developed from research findings is consistent with the data), evaluative validity (i.e., the extent to which an evaluation framework can be applied to the objects of study, as opposed to a descriptive, interpretive, or explanatory one), and generalizability (i.e., the extent to which a researcher can generalize the account of a particular situation, context, or population to other individuals, times, settings, or context). With regard to the last validity type, Maxwell differentiates internal generalizability from external generalizability, with the former referring to the generalizability of a conclusion within the underlying setting or group, and the latter pertaining to generalizability beyond the group, setting, time, or context. According to Maxwell, internal



## THE VALIDITY ISSUE IN MIXED RESEARCH

generalizability is typically more important to qualitative researchers than is external generalizability (see also, Maxwell, 2005).

Onwuegbuzie and Leech (in press-a) conceptualized what they called the *Qualitative Legitimation Model*, which contains 29 elements of legitimation for qualitative research at the following three recursive stages of the research process: research design/data collection, data analysis, and data interpretation.<sup>1</sup> As illustrated in Figure 3, the following threats to internal credibility are viewed as pertinent to qualitative research: ironic legitimation, paralogical legitimation, rhizomatic legitimation, voluptuous (i.e.,

embodied) legitimation, descriptive validity, structural corroboration, theoretical validity, observational bias, researcher bias, reactivity, confirmation bias, illusory correlation, causal error, and effect size. Also in this model, the following threats to external credibility were identified as being pertinent to qualitative research: catalytic validity, communicative validity, action validity, investigation validity, interpretive validity, evaluative validity, consensual validity, population generalizability, ecological generalizability, temporal generalizability, researcher bias, reactivity, order bias, and effect size.

Figure 1. Threats to Internal and External Validity

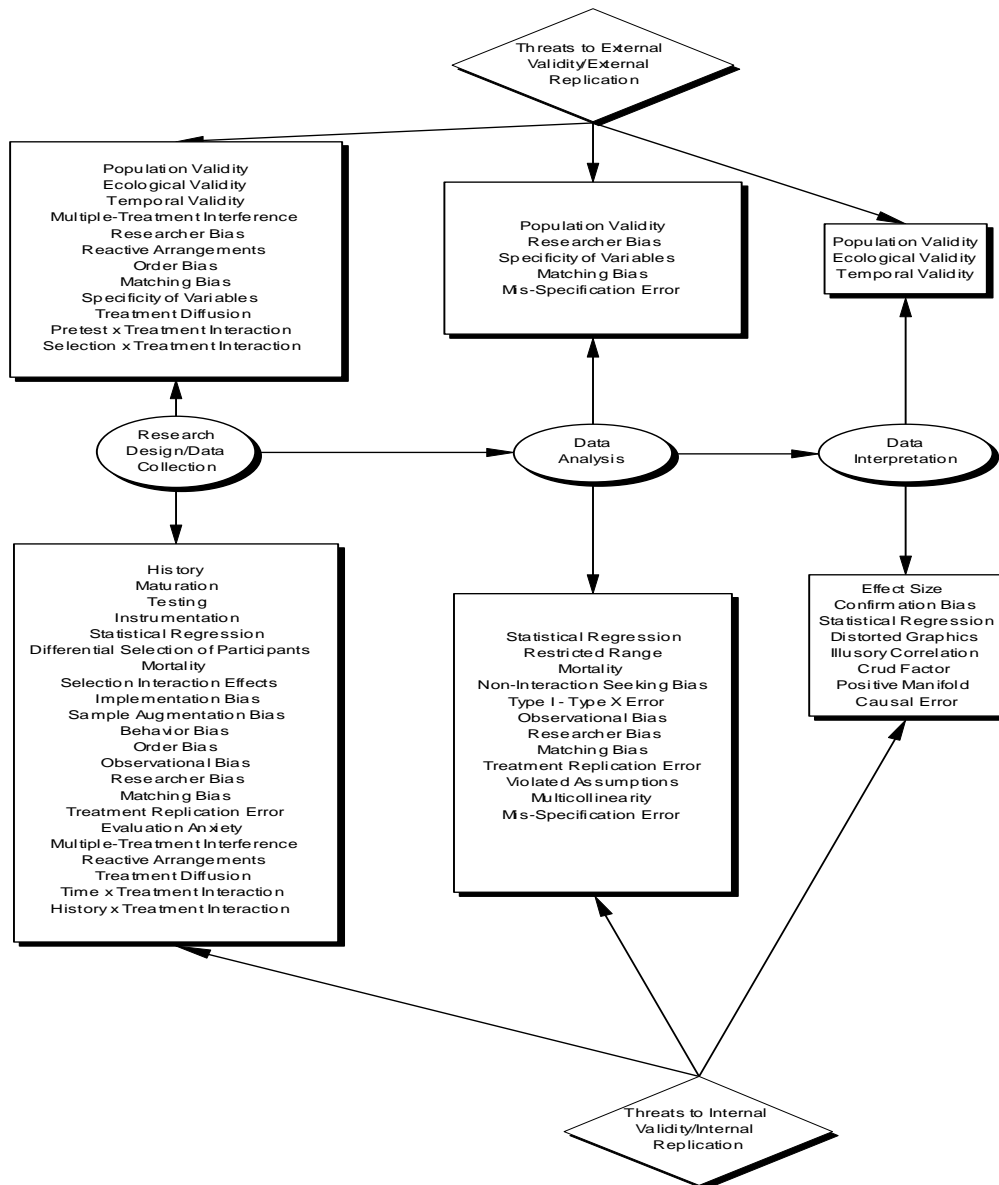
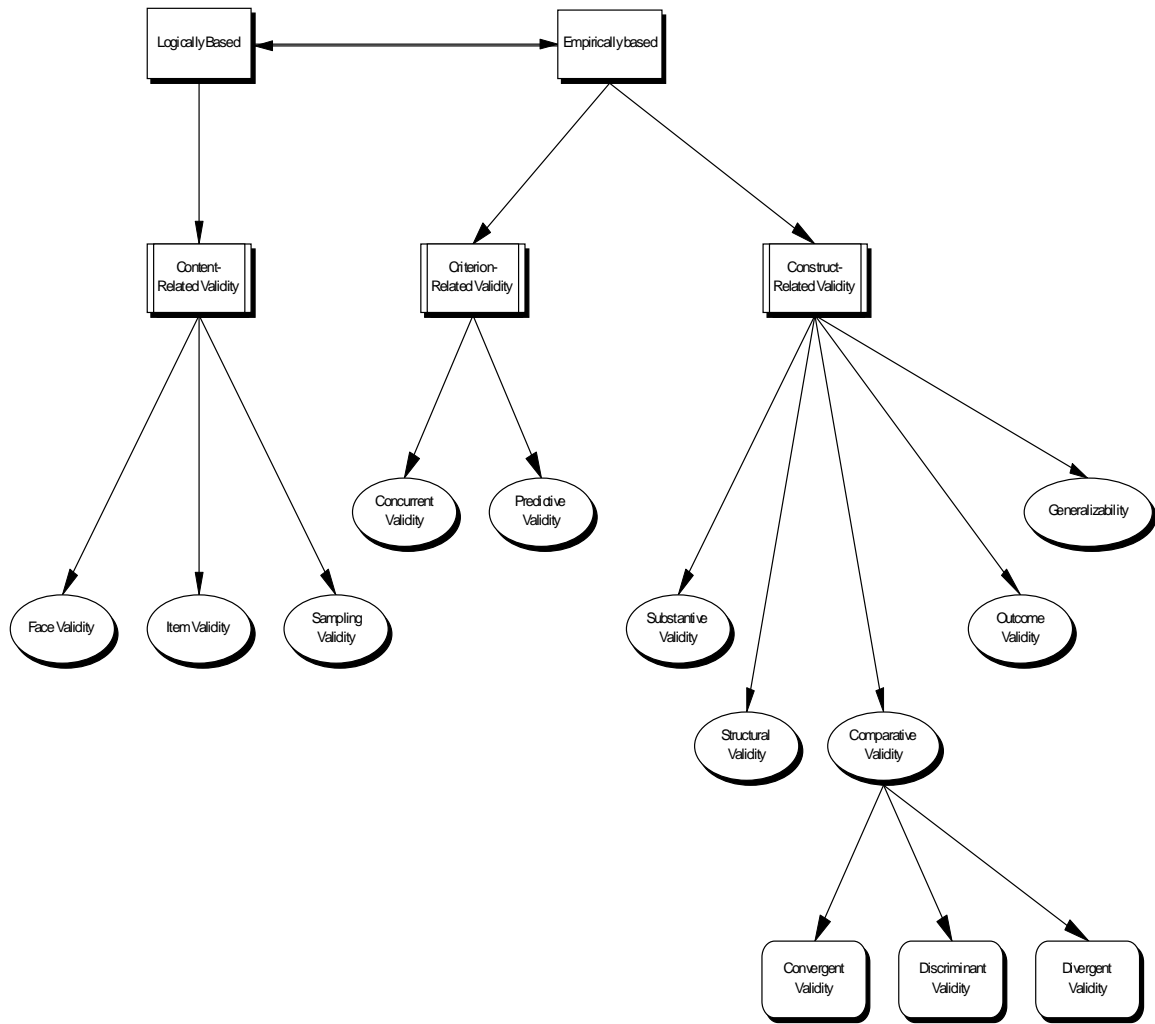


Figure 2. Schematic Representation of Instrument Score Validity



Validity in Mixed Research

Because of the association with the quantitative conceptualization of the research process, the term validity has generally been replaced by the term trustworthiness within qualitative research. The vast and important literature on trustworthiness is exemplified and discussed in the following references from the qualitative research literature: Creswell (1998), Glaser and Strauss (1967), Kvale (1995), Lather (1986, 1993), Lincoln and Guba (1985, 1990), Longino (1995), Maxwell (1992, 1996), Miles and Huberman (1984, 1994), Onwuegbuzie and Leech (in press-a), Schwandt (2001), Strauss and Corbin (1998), and Wolcott (1990).

Mixed research involves the mixing of quantitative and qualitative methods or paradigm characteristics into research studies (Johnson & Onwuegbuzie, 2004; Onwuegbuzie & Johnson, 2004; Tashakkori & Teddlie, 1998, 2003). According to the *fundamental principle of mixed research*, it often should involve the combining of quantitative and qualitative methods, approaches, and concepts that have complementary strengths and nonoverlapping weaknesses (Brewer & Hunter, 1989; Johnson & Turner, 2003). This principle is meant to be viewed broadly; *it is not limited to triangulation or corroboration*. The words “complementary strengths” are meant to include all of the strengths of qualitative

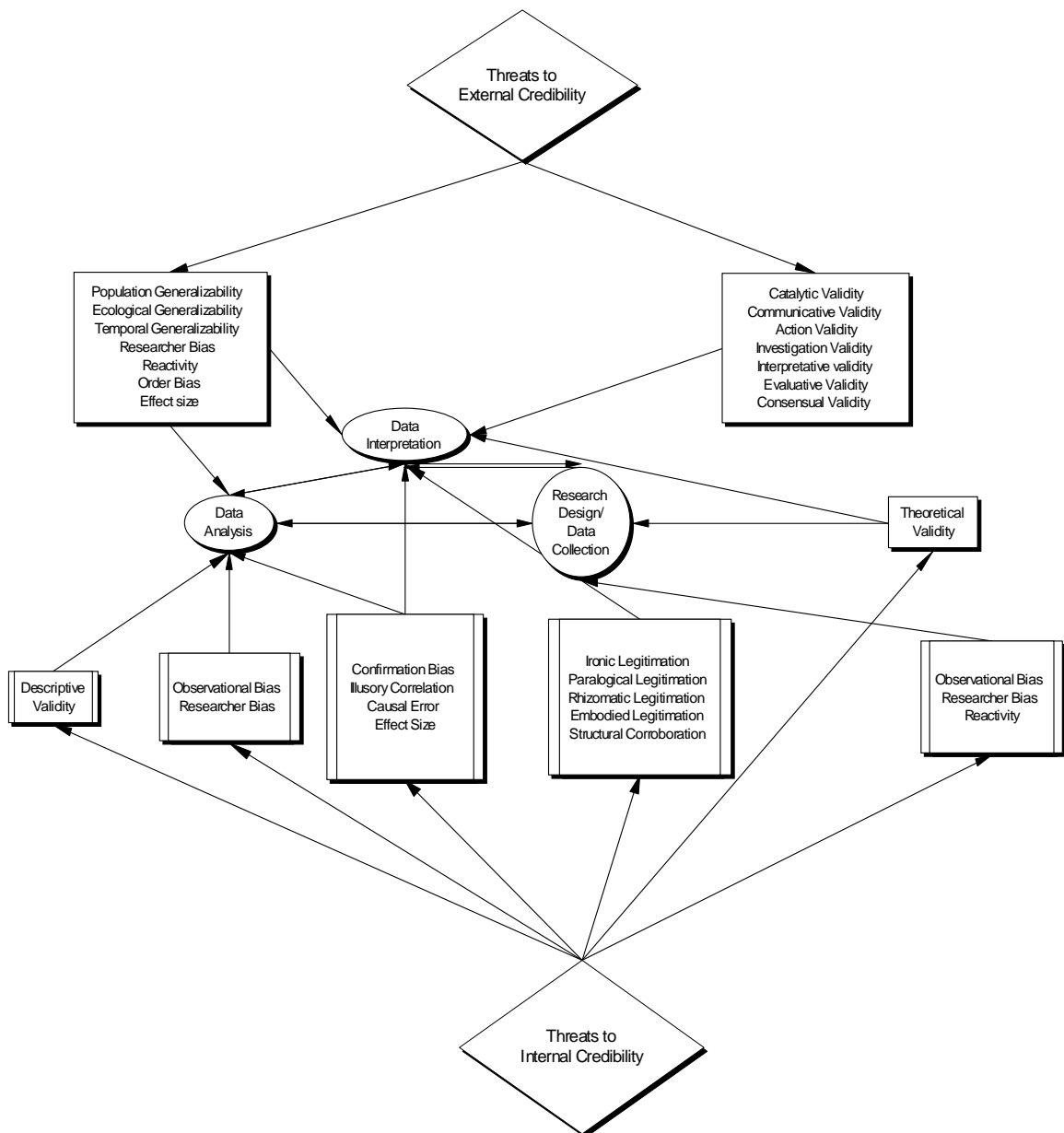
and quantitative research. Therefore, the principle can be used for the five traditional purposes of mixed research identified by Greene, Caracelli, and Graham (1989). By “complementary strengths” we are implying a putting together of different approaches, methods, and strategies in multiple and creative ways.

Mixed research still is plagued by the problems of representation, integration, and legitimation (Onwuegbuzie, in press). The *problem of representation* refers to the difficulty in capturing (i.e., representing) lived experiences using text in general and words and numbers in particular. The *problem of*

*legitimation* refers to the difficulty in obtaining findings and/or making inferences that are credible, trustworthy, dependable, transferable, and/or confirmable. Indeed, in many instances, these problems are exacerbated in mixed research because both the quantitative and qualitative components of studies bring into the setting their own problems of representation and legitimation, likely yielding either an additive or a multiplicative threat—hence the *problem of integration*.

Mixed research can be conceptualized as combining quantitative or qualitative research in a

Figure 3. Qualitative Legitimation Model



concurrent, sequential, conversion (Tashakkori & Teddlie, 2003; Teddlie & Tashakkori, 2006), parallel (Onwuegbuzie & Leech, 2004a), or fully mixed (Leech & Onwuegbuzie, 2005; Teddlie & Tashakkori, 2006) manner. Quantitative and qualitative approaches can be combined in these ways whether the study represents primary research (Johnson & Onwuegbuzie, 2004; Teddlie & Tashakkori, 2006) or a mixed synthesis of the extant literature (i.e., integrating the findings from both quantitative and qualitative studies in a shared area of empirical research; Sandelowski, Voils, & Barroso, 2006). Further, quantitative and qualitative approaches can be combined in these ways regardless of which approach has priority in the study (cf. Creswell, Shope, Plano Clark, & Green, 2006). In basic concurrent mixed designs, the following three conditions hold: (a) both the quantitative and qualitative data are collected separately at approximately the same point in time, (b) neither the quantitative nor qualitative data analysis builds on the other during the data analysis stage, and (c) the results from each type of analysis are not consolidated at the data interpretation stage, until *both* sets of data have been collected and analyzed separately, and (d) after collection and interpretation of data from the quantitative and qualitative components, a meta-inference is drawn which integrates the inferences made from the separate quantitative and qualitative data and findings.

In basic sequential mixed designs, data collected and analyzed from one phase of the study (i.e., quantitative/qualitative data) are used to inform the other phase of the investigation (i.e., qualitative/quantitative data). Here, the data analysis begins *before* all the data are collected. At the highest level of integration, referred to as sequential mixed model studies, “multiple approaches to data collection, analysis, and inference are employed in a sequence of phases. Each phase, by itself, may use a mixed approach and provide conceptual and/or methodological grounds for the next one in the chain” (Tashakkori & Teddlie, 1998, pp. 149-150). Sequential mixed designs also can be applied when conducting what Chen (1990, 2006) conceptualizes as theory-driven evaluations, via the following two strategies: (a) switch strategy (e.g., first applying qualitative methods to illuminate program theory of stakeholders and then use quantitative methods to assess the program theory) and (b) contextual overlaying strategy (e.g., utilizing qualitative approaches to collect contextual information for facilitating the interpretation of quantitative data or reconciling findings).

Conversion mixed designs involve data transformation wherein one data form is converted into the other and then subsequently analyzed (Teddlie & Tashakkori, 2006). That is, the other data type evolves from the original data type either by converting the

data from quantitative to qualitative or from qualitative to quantitative. Moreover, conversion occurs via techniques such as quantizing data (i.e., transforming the qualitative data to a numerical form; Tashakkori & Teddlie, 1998) or qualitzing data (i.e., converting quantitative data into data that can be analyzed qualitatively; Tashakkori & Teddlie, 1998). Both data types are analyzed/re-analyzed, and inferences are made based on both sets of analyses.

In parallel mixed designs, the data are collected and analyzed separately. In this respect it is similar to concurrent designs. However, while inferences are made in concurrent designs on both sources of data in an integrated manner, in parallel mixed designs, each data source leads to its own set of inferences, and no attempt is made to reach what Tashakkori and Teddlie (2003) refer to as a “meta-inference” (p. 686), in which both sets of inferences are combined into a coherent whole. Such designs lead either to (a) two separate reports that would be presented or published separately or (b) two separate write-ups that are presented in two distinct sections of the same report. Whereas some researchers do not consider these designs as representing mixed research (e.g., Yin, 2006) but rather quasi-mixed designs (e.g., Teddlie & Tashakkori, 2006), other researchers do (e.g., Onwuegbuzie & Leech, 2004a). Given the formative stage of mixed research, we see, as does Greene (2006), these current disagreements as being good for the field as it continues to develop through hard conceptual and empirical work.

Fully mixed research designs (Leech & Onwuegbuzie, 2005), also known as fully integrated mixed research designs (Tashakkori & Teddlie, 2003; Teddlie & Tashakkori, 2006), involve mixing quantitative and qualitative approaches in an interactive way at all stages of the investigation (i.e., research objective, type of data/operations, type of analysis/inference; Onwuegbuzie & Johnson, 2004) such that at each stage, one approach (e.g., quantitative) influences the formulation of the other approach (e.g., qualitative). We agree with Teddlie and Tashakkori (2006) that fully mixed (and nearly fully mixed designs) are attractive because of the multiple points of integration.

Because of the complexity involved in combining qualitative and quantitative studies either in a concurrent, sequential, conversion, parallel, or fully mixed manner, mixed research gives rise to what we call the problem of integration. Surrounding this problem is the extent to which combining quantitative and qualitative approaches can address each of Greene et al. (1989) five empirically derived, general purposes of mixed-methodological research studies: (a) triangulation (i.e., seeking convergence and corroboration of findings from different methods that study the same phenomenon); (b) complementarity

(i.e., seeking elaboration, illustration, enhancement, and clarification of the findings from one method with results from the other method); (c) development (i.e., using the findings from one method to help inform the other method); (d) initiation (i.e., discovering paradoxes and contradictions that lead to a re-framing of the research question); and (e) expansion (i.e., seeking to expand the breadth and range of inquiry by using different methods for different inquiry components). More generally, the problem of integration pertains to the extent to which combining quantitative and qualitative research techniques addresses one or more of Collins, Onwuegbuzie, and Sutton's (2006) four rationales for mixing or combining qualitative and quantitative approaches: participant enrichment (i.e., mixing quantitative and qualitative techniques to optimize the sample using techniques that include recruiting participants, engaging in activities such as Institutional Review Board debriefings, and ensuring that each participant selected is appropriate for inclusion), instrument fidelity (e.g., assessing the appropriateness and/or utility of existing instruments; creating new instruments; performance of human instruments), treatment integrity (i.e., assessing fidelity of intervention), and significance enhancement (e.g., facilitating thickness and richness of data; augmenting interpretation and usefulness of findings).

The problem of integration motivates us to ask questions such as the following: Is it misleading to triangulate, consolidate, or compare quantitative findings and inferences stemming from a large random sample on equal grounds with qualitative data arising from a small purposive sample? How much weight should be placed on quantitative data compared to qualitative data? Are quantitatively confirmed findings more important than findings that emerge during a qualitative study component? When findings conflict, what is one to conclude?

Before discussing the issue of integration more carefully, we will point out that we do not believe that the goal of mixed research is to replace either quantitative or qualitative research. Rather, the goal of this third type of research is to utilize the strengths of two or more approaches by combining them in one study, and by attempting to minimize the weaknesses of approaches in mixed designs. Philosophically, mixed research generally follows philosophical and methodological pragmatism (with a very broad and inclusive ontological realism where virtually everything a qualitative or quantitative researcher deems to be real can be considered, in some sense, to be real, including subjective realism, intersubjective realism, and objective realism). (See Johnson & Onwuegbuzie, 2004 for more discussion on the role of pragmatism in mixed research, and Sanders, 1997 for more discussion about inclusive ontology.) By

pragmatism, we mean to search for workable solutions through the practice of research (e.g., follow the fundamental principle of mixed research, including the use of designs and criteria that are situation and context appropriate) to help answer questions that we value and to provide workable improvements in our world (i.e., help in bringing about desired outcomes). Our pragmatism includes a healthy dose of pluralism by which we mean that it is not logically contradictory to claim that quantitative and qualitative research are both useful, even if, at times, they appear to be contradictory; perhaps what is seen as contradictory are different perspectives that are complementary and enable one to more fully to see his or her world. Further, different standards of quality will be useful for different people in different contexts (see Patton's, 2002, five different sets of criteria for judging the quality of qualitative research), which is reasonable as long as one makes these standards clear to avoid arguments based on equivocation (which can produce misunderstandings based on nothing more than different meanings of terms being used by different people because they "talk past" one another). Clarity of language use is especially important when people from different communities are the interlocutors. Arguments about values can be healthy, and at this time more discussion among qualitative, quantitative, and mixed researchers about values needs to take place in order to understand better each other. Different researchers have different values and beliefs about research approaches for addressing important questions, and this diversity when put together is not a problem; we see it as a *potential strength* of research and practice, especially when simple and clear solutions are not readily forthcoming. In short, diversity of this sort is not a problem needing to be fixed by someone. Our pragmatism also is eclectic, by which we are referring to the inclusion of multiple quantitative and qualitative techniques in one's briefcase and then selecting combinations of assumptions, methods, and designs that best fit one's research questions of interest.

In every mixed research study researchers must deal with the problems of representation, legitimation, and integration but discussions about validity issues that characterize these problems are still in relative infancy. Developing justified inferences is at the center of many problems in mixed research. In fact, Teddlie and Tashakkori (2003) and Tashakkori and Teddlie (2003) identified drawing inferences as one of the six unresolved issues and controversies in mixed research. The purpose of the remainder of this paper is to contribute to the present dialogue about validity (trustworthy or defensibility or quality) issues in mixed research. We will discuss the following three unresolved and, at times, contentious issues in the domain of validity in mixed research:

- (a) Labels for criteria for assessing mixed research studies;
- (b) conceptualization of legitimation in mixed research studies; and
- (c) identifying some types of legitimation for mixed research.

#### *Label for Criteria for Assessing Mixed Research Studies*

As noted by Teddlie and Tashakkori (2003), a primary decision that confronts the field of mixed research is what to call the concept of validity in mixed research. Although the term “validity” is routinely used in quantitative research, this term is disliked by many qualitative researchers. In fact, as noted by Schwandt (2001), some qualitative researchers object to the concept of validity based on their rejection of the correspondence theory of truth. They argue that because validity is the test of this correspondence, validity does not exist because there simply is no single reality, with truth being partially arbitrary as individuals interact with their worlds. These researchers also believe in *fallibilism* (i.e., all procedures for establishing legitimation represent “fallible means of making a case for a plausible and credible account”; Schwandt, 2001, pp. 268-269). Some qualitative researchers, although they believe that some validities are useful, contend that validity is always relative to a particular context, situation, language system, or worldview (Schwandt, 2001). These researchers refer to *contextualization* (i.e., legitimation represents the standards set by a particular community at a specific time and place). Some qualitative researchers refute any relationship between validity and objectivism, the latter of which is viewed as foundational. These researchers are referred to as representing *strong or radical relativism* (i.e., no single account can be judged as being superior to any other). Some qualitative researchers (i.e., postmodernists) view the concept of validity (and the word) as representing a debunked modernist perspective that champions universal rationality, rules, order, logic, and the like. Thus, we conclude that use of the word *validity* in mixed research can be counterproductive.

One attractive solution to this problem is for mixed researchers to use an alternative word that is more acceptable to both quantitative and qualitative researchers. This solution involves what Teddlie and Tashakkori (2003) refer to as “using a bilingual nomenclature” (p. 12). In this respect, a possible term that might be acceptable to both quantitative and qualitative investigators is *legitimation*. This would be consistent with its use in the Onwuegbuzie (2003) *Quantitative Legitimation Model* and the Onwuegbuzie and Leech (in press-a) *Qualitative Legitimation Model* presented in Figures 1 and 3. That is not to suggest that quantitative researchers should refrain from using the

term validity or that qualitative researchers should cease using terms such as trustworthiness, credibility, plausibility, and dependability. It only is to suggest that in the context of discussing the overall criteria for assessment of mixed research studies, we recommend that the term legitimation, or a similarly descriptive and inclusive term, be used.

#### *Conceptualization of Legitimation in Mixed Research Studies*

In one of the very few essays written on the topic of validity or quality criteria in mixed research, Teddlie and Tashakkori (2003) stated that mixed methods researchers “should adopt a common nomenclature transcending the separate QUAL and QUAN orientations when the described processes (QUAL and QUAN) are highly similar and when appropriate terminology exists” (p. 12). Because inferences are made in research studies regardless of whether the associated interpretation is inductive or deductive in nature, these authors contended that the concept of “inference” transcends quantitative and qualitative research and they recommended that *inference quality* be used as the mixed research term for validity. This use has much merit, and we attempt to build on it below.

Teddlie and Tashakkori (2003) conceptualized inference quality as being associated with the following two research components: design quality and interpretive rigor. Design quality refers to the standards used for the evaluation of the methodological rigor of the mixed research study, whereas interpretive rigor pertains to the standards for evaluating the validity of conclusions. Teddlie and Tashakkori also presented the term *inference transferability* to denote the generalizability of the findings (for both quantitative and qualitative research), which comprises population transferability (i.e., transferability to other individuals, groups, or entities), ecological transferability (i.e., transferability to other contexts or settings), temporal transferability (i.e., transferability to other time periods), and operational transferability (i.e., transferability to other methods of measuring behaviors). Teddlie and Tashakkori appropriately differentiated data quality from inference quality. What also is appealing about Teddlie and Tashakkori’s conceptualization of inference quality is their identification of the following four (non-exhaustive and not mutually exclusive) criteria for evaluation: (a) within-design consistency (i.e., “consistency of the procedures/design of study and from which the inference emerged”; p. 40); (b) conceptual consistency (i.e., “degree to which the inferences are consistent with each other and with the known state of knowledge and theory”; “consistency of inferences with each other within a study [cross-inference consistency]”; and “consistency of inference with current state of

knowledge and theory [theoretical consistency]"; p. 40); (c) interpretive agreement (or consistency) (i.e., "consistency of interpretations across people"; p. 40); and (d) interpretive distinctiveness (i.e., the "degree to which the inferences are distinctively different from other possible interpretations of the results and rival explanations are ruled out" p. 40).

Building on the work of Teddlie and Tashakkori (2003), Tashakkori and Teddlie (2006) proposed an integrative model of quality that also comprises design quality and interpretive rigor. According to their model, design quality comprises (a) within-design consistency (as defined earlier), (b) design suitability (i.e., whether the methods of the study are appropriate for addressing the research question(s); and the design is consistent with the research question), (c) design fidelity (i.e., whether the procedures are implemented with quality and rigor; the methods are capable of capturing meaning, associations, or effects; and the components of the design such as sampling and data collection procedures, are implemented adequately); and (d) analytic adequacy (i.e., whether the data analysis techniques are appropriate for addressing the research question(s)). Interpretive rigor consists of (a) interpretive agreement (as defined earlier), (b) interpretive distinctiveness (as defined earlier), (c) interpretive consistency (i.e., whether the inferences closely follow the relevant findings in terms of type, intensity, and scope; and the multiple inferences made on the basis of the findings are consistent with each other), (c) theoretical consistency (i.e., whether the inferences are consistent with theory and the state of knowledge in the field), and integrative efficacy (i.e., whether the meta-inference adequately incorporates the inferences stemming from quantitative and qualitative phases of the study).

Teddlie and Tashakkori's (2003) and Tashakkori and Teddlie's (2006) conceptualizations present inference as an outcome. However, as appealing and useful as their conceptualization is, we believe it needs some elaboration and extension. We see useful extensions of their model in two ways. First, we view legitimation as a process, not just an outcome. Indeed, we believe that legitimation checks should occur at each stage of the mixed research process. Thus, the Quantitative Legitimation Model and Qualitative Legitimation Model, shown earlier, can be used for assessing legitimation of the quantitative and qualitative components of the study, respectively. While, clearly, making inferences is a vital part of the research process, giving inference quality primary emphasis could give the false impression that one does not have to scrutinize as carefully some of the other steps of the research process. Also, it is not clear yet what role the validity types presented in this paper (and in the selected references) will play in the evaluation process. Moreover, legitimation in mixed research

should be seen as a continuous process rather than as a fixed attribute of a specific research study. Mixed research tends to be iterative and interactive (Onwuegbuzie & Johnson, 2004) such that, in a sense, *inference closure* (i.e., being able to make definitive statements about the quality of inferences made) might never be fully reached within a particular study or even over a series of systematically linked studies. We look forward to future dialogue about these issues as we all try to advance the field forward.

#### *Some Types of Legitimation for Mixed Research*

As noted earlier, the problems of representation and integration in mixed research suggest the need to identify specific legitimation issues that are not associated with monomethod designs. However, these legitimation issues are not addressed fully in Teddlie and Tashakkori's (2003) and Tashakkori and Teddlie's (2006) inference quality frameworks, nor do these issues appear to have been addressed, to date, in any other framework. Thus, we now will outline a new typology of legitimation types in mixed research for consideration, dialog, and refinement. Our typology currently is in its infancy, and it contains nine legitimation types. These legitimation types are summarized in Table 1. Each of these types of legitimation is discussed next.

*Sample integration legitimation.* This legitimation type applies to situations in which a researcher wants to make statistical generalizations from the sample participants to a larger target population. Unless exactly the same individuals or groups are involved in both the qualitative and quantitative components of a study, constructing meta-inferences by pulling together the inferences from the qualitative and quantitative phases can be problematic. For example, a researcher might conduct a concurrent design in which inferences made from quantitative data yielded by a large random sample were integrated (i.e., into a meta-inference) with inferences made from qualitative data arising from a smaller subset of this sample or from an entirely different group of people. However, if this meta-inference was generalized to the underlying population from which the large random sample was selected, it may not be justified for this inference to include inferences from the qualitative component, especially if the associated subsample is very small or if it forms a separate group of people. That is, because of the unrepresentative sample from the qualitative phase, the ensuing meta-inference might be poor (statistically speaking), which, in turn, would affect statistical generalizability (i.e., population transferability). To the degree to which the qualitative participants are similar a quantitative random sample, the problem will be reduced.

Both the inference quality and generalizability are even poorer if the quantitative sample is nonrandom, as

is the case in the vast majority of empirical research studies (Onwuegbuzie & Leech, 2004b), and/or small. Even if the qualitative sample represented a random subset of the quantitative sample, as might be the case in a sequential mixed design, the meta-inference quality might still be poor. As such, when the researcher's goal is to make a statistical generalization we would urge caution in considering Teddlie and Tashakkori's (2003) contention that "in evaluating the quality of inferences in mixed research, the issue of dominance or priority of one methodological approach (e.g., QUAL-*quan*, *qual*-QUAN) over another is not very important" (p. 41). Specifically, the use of a dominant-less dominant design is more likely to lead to

the combining of a strong inference (dominant phase) with a weak inference (less dominant design). If the inferences stemming from the quantitative and qualitative phases were consistent, then the meta-inference quality likely would be higher. However, a mixed methods researcher should not assume that this will always be the case. Regardless, criteria are needed to be developed to identify the range of conditions under which combining inferences from the qualitative and quantitative components of a study leads to meta-inference quality. Indeed, as noted by Collins, Onwuegbuzie, and Jiao (in press) and Onwuegbuzie and Collins (in press), the relationship between the quantitative and qualitative sampling designs (i.e.,

Table 1  
Typology of Mixed Methods Legitimation Types

Legitimation Type	Description
Sample Integration	The extent to which the relationship between the quantitative and qualitative sampling designs yields quality meta-inferences.
Inside-Outside	The extent to which the researcher accurately presents and appropriately utilizes the insider's view and the observer's views for purposes such as description and explanation.
Weakness Minimization	The extent to which the weakness from one approach is compensated by the strengths from the other approach.
Sequential	The extent to which one has minimized the potential problem wherein the meta-inferences could be affected by reversing the sequence of the quantitative and qualitative phases.
Conversion	The extent to which the quantizing or qualizing yields quality meta-inferences.
Paradigmatic mixing	The extent to which the researcher's epistemological, ontological, axiological, methodological, and rhetorical beliefs that underlie the quantitative and qualitative approaches are successfully (a) combined or (b) blended into a usable package.
Commensurability	The extent to which the meta-inferences made reflect a mixed worldview based on the cognitive process of Gestalt switching and integration.
Multiple Validities	The extent to which addressing legitimation of the quantitative and qualitative components of the study result from the use of quantitative, qualitative, <i>and</i> mixed validity types, yielding high quality meta-inferences.
Political	The extent to which the consumers of mixed methods research value the meta-inferences stemming from <i>both</i> the quantitative and qualitative components of a study.



sampling scheme, sample size) is crucial to assessing meta-inference quality. Additional considerations regarding sample quality also should be considered when examining this type of legitimation. For example, in a qualitative sample, sometimes saturation is a useful criterion with regard to the conclusions (Guest, Bunce, & Johnson, 2006; Onwuegbuzie & Leech, 2005, in press-b; Strauss & Corbin, 1998; Teddlie & Yu, 2006). Sometimes theoretical generalizations can be made even in the absence of statistical sampling methods; for example, Yin (1994) has demonstrated this with some of the classic sociological community studies conducted in the twentieth century. In sum, it is essential that the way individuals and groups are selected be considered, and that additional consideration be made on how to combine legitimately different sets of people for use in making quality meta-inferences.

*Inside-outside legitimation.* As noted by Currall and Towler (2003), “etic refers to the trained observer’s analysis of ‘raw’ data, whereas emic refers to how those data are interpreted by an ‘insider’ to the system or organization (Pike, 1967)” (p. 522). In other words, the *emic viewpoint* is the viewpoint of the group member, the insider. The *etic viewpoint* is that of the “objective” outsider looking at and studying the group. One can even speak of emic terms (language used by the group members) and etic terms (the language used by the outsider researcher) (Johnson & Christensen, 2004). Along the same lines as sample integration legitimation, when making meta-inferences by combining inferences from the qualitative and quantitative phases of a study, there are times when researchers should assess insider-outsider legitimation. This refers to the degree to which the researcher accurately presents and utilizes the insider’s view *and* the observer’s view. The ability to do this can be compromised when a researcher is ethnocentric or, on the other hand, when a researcher becomes so involved with the group that he or she “goes native.”

A strategy for obtaining a justified etic viewpoint is for the researcher to use peer review; that is, the research can have another (disinterested and trained in social research) outsider/researcher examine the interpretations being made, the conceptualizations, and the relationship between the data and the conclusions. An important strategy for obtaining a justified insider viewpoint is member checking or participant review (i.e., have group members or participants assess the researcher’s interpretations). A strategy for obtaining a justified meta-inference typically will be for everyone on the research team as well as some researchers outside of the team and participants inside the group under study to review the data and integration. In other words, the researcher should seek insider-outsider legitimation for the qualitative part of a study, for the quantitative part, and when the parts are put together or

integrated (e.g., by maintaining a well informed and balanced perspective when collecting, analyzing, and interpreting what the whole set of qualitative and quantitative data mean). One might be able to make the case that quantitative research often seeks the objective outsider view, that qualitative research often seeks the insider’s view, and that mixed research seeks to balance fully these two viewpoints.

*Weakness minimization legitimation.* Mixed research is in the optimal position for maximizing this form of legitimation simply because the researcher is able systematically to design a study that combines two or more methods. The key, however, is that the researcher must consciously and carefully assess the extent to which the weakness from one approach can be compensated by the strengths from the other approach and then plan and design the study to fulfill this potential; the researcher also must use this knowledge when combining, weighting, and interpreting the results. We refer to this process as *weakness minimization legitimation*. The greater the extent that the weakness from one approach is compensated by the strengths from the other approach, the more likely that combining a weak inference with a strong inference will lead to a superior or high quality meta-inference.

*Sequential legitimation.* When a sequential mixed research design is used, it is possible that the meta-inference that arises is solely or largely the effect of the sequencing itself. For example, if the results and interpretations would have been different if the order the quantitative and qualitative phases originally presented had been reversed, then this would indicate that the sequencing itself was a threat to legitimation. One method of assessing this is by changing the sequential design to a multiple wave design, in which the quantitative and qualitative data collection and data analysis phases oscillate multiple times (Sandelowski, 2003).

*Conversion legitimation.* All inferences or meta-inferences that are made after quantizing and/or quantizing the data must be scrutinized. The extent to which these data conversion techniques lead to interpretable data and high inference quality is called conversion legitimation. For example, a popularized way of quantizing data is by counting. Obtaining counts of the themes present in qualitative data can prevent researchers from over-weighting or under-weighting emergent themes (Sandelowski, 2001). Also, qualitative researchers can sometimes obtain more meaning by obtaining counts of observations in addition to their narrative descriptions (Johnson & Christensen, 2004; Onwuegbuzie & Leech, 2004a; Onwuegbuzie & Teddlie, 2003; Sandelowski, 2001) because counting can provide additional useful information about how often or how many or how much. However, counting is not appropriate for some

types of qualitative data and contexts. As noted by Sandelowski (2001), researchers should avoid the problems associated with verbal counting, misleading counting, over-counting, and acontextual counting. Such problems would affect the meta-inference quality. Similarly, a common method of qualitzing data is via narrative profile formation (i.e., modal profiles, average profiles, holistic profiles, comparative profiles, normative profiles). Such profiles involve constructing narrative descriptions from quantitative data. However, these descriptions can represent an over-generalization of the observed numeric data. Further, it is possible that a profile that emerges from qualitzing (e.g., via average profiles) yields a representation of people that is unrealistic.

*Paradigmatic mixing legitimation.* Combining quantitative and qualitative approaches is sometimes considered to be tenuous because of competing dualisms: epistemological (e.g., objectivist vs. subjectivist), ontological (e.g., single reality vs. multiple reality), axiological (e.g., value free vs. value-bound), methodological (e.g., deductive logic vs. inductive logic), and rhetorical (e.g., formal vs. informal writing style) beliefs. One solution is to use both viewpoints in a study (e.g., have a pure qualitative part and a pure quantitative part each based on the pure assumptions), and then attempt to make meaning from consideration of the two pure components of the study. Another solution is to think in terms of continua rather than dualisms and then take more moderate positions on each continuum: ontological (recognition of multiple affordances, levels of analysis, and disciplinary perspectives about what is studied; recognizing subjective, intersubjective, and objective types of reality; recognizing internal reality, external reality, and most importantly the interaction between the two), epistemological (intersubjective approach to knowledge generation), axiological (distinguishing between internal and external values, admitting and describing the value ladenness of the research; stating one's use of values in setting standards, determining what outcomes are to be valued, interpreting the data, making recommendations, and making explicit how one judged one's own study), and rhetorical (e.g., use of formal and informal writing styles using both impersonal and personal voices). When making meta-inferences, there are times when a researcher should evaluate the extent to which her or his epistemological, ontological, axiological, methodological, and rhetorical beliefs that underlie the quantitative and qualitative approaches are treated as separate *but* complementary or are used in less extreme forms and treated as being compatible. Legitimation comes from the researcher making the use of paradigm assumptions explicit and conducting research that fits with the stated assumptions.

*Commensurability legitimation.* This type of legitimation is based on a *rejection* of Kuhn's and Quine's (and others') concept of incommensurability of findings, theories, language, and worldviews. In order to meet this type of legitimation, the mixed researcher must learn to make Gestalt switches from qualitative lens to a quantitative lens, going back and forth, again and again. We believe this is possible through cognitive and empathy training. (If one believes this is not possible, then one can ignore "commensurability legitimation.") Through an iterative process, a third viewpoint is created, a viewpoint that is informed by, is separate from, and goes beyond what is provided by either a pure qualitative viewpoint or a pure quantitative viewpoint. To the extent that the researcher is able to negotiate cognitively this important Gestalt switch, the meta-inferences will provide a more fully mixed worldview; it will go beyond the provision of both traditional viewpoints by offering a third, well-informed viewpoint based on consideration of both qualitative and quantitative thinking. This argument takes seriously what has been called the *compatibility thesis* (Howe, 1988; Reichardt & Rallis, 1994).

*Multiple validities legitimation.* This legitimation type, which is pertinent in virtually every mixed research study, refers to the extent to which all relevant research strategies are utilized and the research can be considered high on the multiple relevant "validities." For example, when addressing legitimation of the quantitative component, the relevant quantitative validities are addressed and achieved; when addressing legitimation of the qualitative component, the relevant qualitative "validities" are addressed and achieved; and during integration and to allow strong meta-inferences, the relevant mixed legitimation types are addressed and achieved. Relatedly, one should ask to what extent is the whole (i.e., meta-inference quality) greater than the sum of its parts (i.e., inferences arising from each component)?

*Political legitimation.* Onwuegbuzie (in press) has identified four challenges that researchers face when undertaking mixed methods research. One of these challenges is the challenge of politics. This challenge refers to power and value tensions that come to the fore as a result of combining quantitative and qualitative approaches. These tensions include any value or ideologically based conflicts that occur when different researchers are used for the quantitative and qualitative phases of a study, as well as differences in perspectives about contradictions and paradoxes that arise when the quantitative and qualitative findings are compared and contrasted. The challenge of politics also includes the difficulty in persuading the consumers of mixed methods research, including stakeholders and policymakers, to value the meta-inferences stemming from *both* the quantitative and qualitative components

of a study. In traditional quantitative research, decision making and power over the research process is fully in the hands of the centralized researcher in a top down manner. In postmodern qualitative research, much power is placed in the research participants themselves, and the researcher takes the role of collaborator and facilitator. In mixed research, the researcher or research team sometimes will take multiple roles; consequently, mixed researchers will need to deal with issues surrounding multiple or distributed power in the planning, conduct, and the use of research (Fetterman, 2000). A strategy for achieving this form of legitimation is to advocate pluralism of perspectives and to strive to generate practical theory or results that consumers naturally will value because the results answer important questions and help provide workable solutions.

#### Summary and Conclusions

The purpose of this paper has been to contribute to the present dialogue about validity issues in mixed research. We first overviewed the ways validity is viewed and defined in quantitative and qualitative research, and we pointed out that there has been a problem of legitimation in both of these paradigms. Second, we contended that there is also a *problem of representation* and *problem of legitimation* in mixed research. We argued that because mixed research involves combining complementary strengths and nonoverlapping weaknesses of quantitative and qualitative research methods, assessing the validity of findings can be particularly complex—yielding a *problem of integration*. We recommended that validity in mixed research be termed legitimation in order to use a bilingual nomenclature that can be used by both quantitative and qualitative researchers. We briefly summarized Teddlie and Tashakkori's (2003) and Tashakkori and Teddlie's (2006) interesting, emerging evaluation criteria frameworks involving the concept of inference quality. We identified nine new types of legitimation that come to the fore as a result of combining inferences from the quantitative and qualitative components of a mixed research study to form meta-inferences. These nine types of legitimation were sample integration legitimation, insider-outsider legitimation, weakness minimization legitimation, sequential legitimation, conversion legitimation, paradigmatic mixing legitimation, commensurability legitimation, multiple validities legitimation, and political legitimation. These types of legitimation need to be studied more closely in order to determine when and how they operate and how they can be maximized or made to occur. Mixed methods researchers should keep in mind that legitimation represents a process that is analytical, social, aesthetic, emic, etic, political, and ethical, and which must involve the community of

quantitative and qualitative scholars alike who are committed to addressing the multiple problems that can occur in mixed research. This is the only way that the promise of mixed research can be realized in research practice.

We are grateful to Dr. Abbas Tashakkori for his constructive feedback on earlier versions of this manuscript.

#### References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing* (rev. ed.). Washington: American Educational Research Association.
- Bracht, G. H., & Glass, G. V. (1968). The external validity of experiments. *American Educational Research Journal*, 5, 437-474.
- Brewer, J., & Hunter, A. (1989). *Multimethod research: A synthesis of styles*. Newbury Park, CA: Sage.
- 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.
- Chen, H. T. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.
- Chen, H. T. (2006). A theory-driven evaluation perspective on mixed methods research. *Research in the Schools*, 13(1), 75-83.
- Collins, K. M. T., Onwuegbuzie, A. J., & Jiao, Q. G. (in press). Prevalence of mixed methods sampling designs in social science research. *Evaluation and Research in Education*.
- 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.
- 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., Shope, R., Plano Clark, V. L., & Green, D. O. (2006). How interpretive qualitative research extends mixed methods research. *Research in the Schools*, 13(1), 1-11.

- Currall, S. C., & Towler, A. J. (2003). Research methods in management and organizational research: Toward integration of qualitative and quantitative techniques. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 513-526). Thousand Oaks, CA: Sage.
- Denzin, N. K., & Lincoln, Y. S. (2005). The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (3rd ed; pp. 1-32). Thousand Oaks, CA: Sage.
- Fetterman, D. M. (2000). *Empowerment evaluation*. Newbury Park, CA: Sage.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.
- Grenne, J. C. (2006). Toward a methodology of mixed methods social inquiry. *Research in the Schools, 13*(1), 93-98.
- 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.
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(1), 59-82.
- Howe, K. R. (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher, 17*(8), 10-16.
- Johnson, B., & Christensen, L. (2004). *Educational research: Quantitative, qualitative, and mixed approaches* (2nd ed.). Needham Heights, MA: 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, B., & Turner, L. A. (2003). Data collection strategies in mixed methods research. In A. Tashakkori, and C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 297-319). Thousand Oaks, CA: Sage.
- Kvale, S. (1995). The social construction of validity. *Qualitative Inquiry, 1*, 19-40.
- Lather, P. (1986). Issues of validity in openly ideological research: Between a rock and a soft place. *Interchange, 17*, 63-84.
- Lather, P. (1993). Fertile obsession: Validity after poststructuralism. *Sociological Quarterly, 34*, 673-693.
- Leech, N. L., & Onwuegbuzie, A. J. (2005, April). A typology of mixed methods research designs. Invited James E. McLean Outstanding Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.
- Lewin, K. (1952). *Field theory in social science: Selected theoretical papers by Kurt Lewin*. London: Tavistock.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Lincoln, Y. S., & Guba, E. G. (1990). Judging the quality of case study reports. *International Journal of Qualitative Studies in Education, 3*, 53-59.
- Longino, H. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton, NJ: Princeton University Press.
- Longino, H. (1995). Gender, politics, and the theoretical virtues. *Synthese, 104*, 383-397.
- Maxwell, J. A. (1992). Understanding and validity in qualitative research. *Harvard Educational Review, 62*, 279-299.
- Maxwell, J. A. (1996). *Qualitative research design*. Newbury Park, CA: Sage.
- Maxwell, J. A. (2005). *Qualitative research design: An interactive approach* (2nd. ed.). Newbury Park, CA: Sage.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). Old Tappan, N.J.: 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.
- Miles, M. B., & Huberman, A. M. (1984). *Qualitative data analysis: A sourcebook of new methods*. Beverly Hills, CA: Sage.
- Miles, M., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- Onwuegbuzie, A. J. (2003). Expanding the framework of internal and external validity in quantitative research. *Research in the Schools, 10*(1), 71-90.
- Onwuegbuzie, A. J. (in press). Mixed methods research in sociology and beyond. In G. Ritzer (Ed.), *Encyclopedia of sociology*. Cambridge, MA: Blackwell Publishers Ltd.
- 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., & Collins, K. M. T. (in press). A meta-validation model for assessing the score-validity of student teaching evaluations. *Quality & Quantity: International Journal of Methodology*.

- Onwuegbuzie, A. J., & Johnson, R. B. (2004). Mixed method and mixed model research. In B. Johnson & L. Christensen, *Educational research: Quantitative, qualitative, and mixed approaches* (pp. 408-431). Boston, MA: Allyn and Bacon.
- Onwuegbuzie, A. J., & Leech, N. L. (2004a). Enhancing the interpretation of "significant" findings: The role of mixed methods research. *The Qualitative Report*, 9(4), 770-792. Retrieved April 19, 2005, from <http://www.nova.edu/ssss/QR/QR9-4/Onwuegbuzie.pdf>
- Onwuegbuzie, A. J., & Leech, N. L. (2004b). Post-hoc power: A concept whose time has come. *Understanding Statistics*, 3, 151-180.
- Onwuegbuzie, A. J., & Leech, N. L. (2005). The role of sampling in qualitative research. *Academic Exchange Quarterly*, 9, 280-284.
- Onwuegbuzie, A. J., & Leech, N. L. (in press-a). Validity and qualitative research: An oxymoron? *Quality & Quantity: International Journal of Methodology*.
- Onwuegbuzie, A. J., & Leech, N. L. (in press-b). A call for qualitative power analyses: Considerations in qualitative research. *Quality & Quantity: International Journal of Methodology*.
- 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.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.
- Pike, K. L. (1967). *Language in relation to a unified theory of the structure of human behavior*. The Hague, Netherlands: Mouton.
- Reichardt, S. S., & Rallis, S. F. (1994). Qualitative and quantitative inquiries are not incompatible: a call for a new partnership. In C. S. Reichardt & S. F. Rallis (Eds.), *The qualitative-quantitative debate: New perspectives* (pp. 85-91). San Francisco, CA: Jossey-Bass.
- Sandelowski, M. (2001). Real qualitative researchers don't count: The use of numbers in qualitative research. *Research in Nursing & Health*, 24, 230-240.
- Sandelowski, M. (2003). Tables or Tableaux? The challenges of writing and reading mixed methods studies. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 321-350). Thousand Oaks, CA: Sage.
- Sandelowski, M., Voils, C. I., & Barroso, J. (2006). Defining and designing mixed research synthesis studies. *Research in the Schools*, 13(1), 29-40.
- Sanders, J. T. (1997). An ontology of affordances. *Ecological Psychology*, 9(1), 97-112.
- Schwandt, T. A. (2001). *Dictionary of qualitative inquiry* (2nd ed.). Thousand Oaks, CA: Sage.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- 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. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, 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. (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 and behavioral research* (pp. 671-701). Thousand Oaks, CA: Sage.
- Tashakkori, A., & Teddlie, C. (2006, April). *Validity issues in mixed methods research: Calling for an integrative framework*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- 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., & Tashakkori, A. (2006). A general typology of research designs featuring mixed methods. *Research in the Schools*, 13(1), 12-28.
- Teddlie, C., & Yu, F. (2006, April). *Mixed methods sampling procedures: Some prototypes with examples*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

- Wolcott, H. F. (1990). On seeking--and rejecting--validity in qualitative research. In E. W. Eisner & A. Peshkin (Eds.), *Qualitative inquiry in education: The continuing debate* (pp. 121-152). New York: Columbia University, Teachers College Press.
- Yin, R. K. (1994). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.
- Yin, R. K. (2006). Mixed methods research: Are the methods genuinely integrated or merely parallel? *Research in the Schools*, 13(1), 41-47.

Notes

<sup>1</sup> Onwuegbuzie and Leech (2006) note that, unlike the case for quantitative research, in qualitative research, the research design/data collection, data analysis, and data interpretation stages are iterative. That is, in qualitative studies, the research design/data collection, data analysis, and data interpretation stages are recursive, and, thus, non-linear in nature.

<sup>2</sup> Denzin and Lincoln (2000) refer to this as the *crisis of representation*.

<sup>3</sup> Denzin and Lincoln (2000) refer to this as the *crisis of legitimation*.

<sup>4</sup> According to Denzin and Lincoln (2000), there is also a *crisis of praxis* in qualitative research. This crisis asks, "Is it possible to effect change in the world if society is only and always a text?" (p. 17).

## The Contribution of Computer Software to Integrating Qualitative and Quantitative Data and Analyses

Pat Bazeley

*Research Support P/L and Australian Catholic University*

*In published mixed methods studies, qualitative and quantitative approaches have typically been combined by using them side-by-side or sequentially, until the point when the separately generated results are interpreted and conclusions drawn. Integration of different forms of data during analysis, or of different approaches within a single analysis, is much less commonly reported. In this paper, integration of these types is shown to be facilitated by use of computer software. Such integration is seen as occurring: (a) when text and numeric data are combined in an analysis; (b) when data are converted from one form to another during analysis; or (c) when combination and conversion occur together iteratively or in generating blended data for further analyses. Examples are provided to illustrate these various, computer-facilitated approaches to mixing methods.*

It has been argued that “multiple research methods and tools of inquiry—qualitative, non-experimental, and experimental—are essential arsenal for researchers who attempt studies on ‘what works’ in education. Without effective use of a variety of research methods at appropriate times, the quality of evidence on a program suffers, and interpretations of causality are limited” (Chatterji, 2004, p.9). The combination of multiple methods<sup>1</sup> “has a long standing history” in evaluation research where both formative and summative aspects of programs are considered (Rallis & Rossman, 2003; Weiss, 1972). Indeed, “most real-world evaluations pose multiple and diverse questions that cross paradigmatic boundaries, so evaluators tend to be pragmatic in drawing on methods” (Rallis & Rossman, 2003, p.493). Mixing of methods, particularly at the stage of data analysis, has a lesser history, however, perhaps in part because of lack of tools to undertake all but the simplest forms of it.

There is no single approach to undertaking a mixed method study. Those who have attempted typologies have variously arrived at 4, 5, 6 or 8 types of study in which elements of quantitative and qualitative approaches are combined into a unique design (e.g., Creswell, 2003; Greene, Caracelli, & Graham, 1989; Morgan, 1998; Niglas, 2004; Tashakkori & Teddlie, 1998). Johnson and Onwuegbuzie (2004) outline the basis for even more elaborate typologies, but conclude by noting that the

design possibilities for combination cannot be thus limited: choices are guided necessarily by the pragmatic demands of the research question, with studies therefore fitting an almost unlimited number of possible designs. These authors then focus (I think more usefully) on the stages one might go through in the process of designing, conducting, and analyzing the data from a mixed methods study. Bryman (2006) critiques the typology approach more generally from the point of view that they are largely built on theoretical modeling, rather than a review of research in practice (the exceptions being those by himself, Greene et al., 1989, and Niglas, 2004).

### Integration in Mixed Methods Research

One of the critical decision points, and a way in which mixed methods studies might be differentiated, is the point at which elements of quantitative and qualitative approaches are brought together (i.e., integrated), whether that be in the design of the question, at data collection, data analysis, at the point of interpretation, or some combination of these (Caracelli & Greene, 1993; Creswell, 2003). Most commonly, integration of approaches occurs only, or primarily, at the point of final interpretation for the study (Bryman, 2006; Greene et al., 1989); that is, results from quantitative and qualitative components of a study are considered in relation to each other primarily as conclusions are being drawn.

Bryman (2006) found the majority (57%) of the 232 social science articles he reviewed used a combination of a separately administered survey

---

Correspondence for this article should be addressed to Pat Bazeley, Research Support, P.O. Box 2005, Bowral NSW 2576 Australia. Email: pat@researchsupport.com.au

## THE CONTRIBUTION OF COMPUTER SOFTWARE TO INTEGRATING QUALITATIVE AND QUANTITATIVE DATA AND ANALYSES

instrument and qualitative interviewing (mostly in a cross-sectional design), whereas in approximately 27% both quantitative and qualitative data were derived from a single data source (the majority of these being a survey which included open ended questions). Indeed, some have argued for total separation of the qualitative and quantitative components of a multimethod study, with integration considered legitimate *only* at the point of final interpretation (e.g., Morse, 2003; Sale, Lohfield, & Brazil, 2002). The purpose of using multiple methods in studies where quantitative and qualitative data are treated separately is generally to attempt to validate the findings by having corroborative evidence derived from different methods (classically referred to as methodological triangulation), or more often, to explain or complement findings from one method by using another (Bryman, 2006; Greene et al., 1989). Thus, for example, the findings of a quantitative study might be ‘fleshed out’ with qualitative data, or the different sources might contribute different aspects to build a more complete picture. These approaches do not pose a particular or new challenge with regard to analytic procedures as the researcher employs standard statistical and text analysis procedures as appropriate to each separate set of data.

Relatively few studies, even among those using mixed methods, report integration at the stage of data analysis: Greene et al. (1989) found 5 only in their sample of 57 evaluation studies and Bryman also noted, when presenting a preliminary report of his 2006 paper,<sup>2</sup> that just 7 of the 232 studies reviewed used an approach involving transformed data. Niglas (2004), in contrast, reported a much higher proportion, classifying more than 50% of the 145 mixed methods studies she identified within her sample of 1,156 educational articles as having integrated data analysis. The difference lies in the definition of what makes for integrated data analysis: Niglas included any study in this category that made a numeric report from qualitative data, such as indicating the number or proportion of people interviewed who mentioned a particular theme or issue. She notes that “real integration of qualitative and quantitative approaches” before the discussion was “very rare” (personal communication, February 1, 2006).

### Strategies for Integration

Caracelli and Greene (1993) identified four integrative strategies for mixed methods analysis: (a) data transformation, in which one form of data are transformed into another for further analysis; (b) typology development, in which a classification of concepts or categories developed from one set of data is applied to another; (c) extreme case analysis, in which the outliers or residuals revealed by one analysis

are explored using alternative data or methods; and (d) data consolidation/merging to create new variables for use in further analysis. Iterative application of different analysis strategies was seen to have value in further explicating the initial analyses of either or both sources. Indeed, integration of mixed-form analyses was most evident when data from one type was used in analyses of the other type, with the intent of reapplying the results to further the analysis of either data type.

The mixed methods research purpose most frequently served by integration of analyses is initiation, that is, to be provocative and bring fresh perspectives through contradiction and (intended or unintended) discovery of paradox (Caracelli & Greene, 1993; Greene et al., 1989; Rossman & Wilson, 1985). Caracelli and Greene note, however, that particular strategies for integration might be used fruitfully also in the context of expansion, development, and complementarity, but that integration is inconsistent with triangulation (defined as corroboration or validation), given the latter requires independence of methods.

Given the potential for enriched understanding that an integrative strategy holds, Caracelli and Greene (1993) ask why integration before interpretation and discussion is so uncommon. Salient suggestions included the impact of the paradigm debates coupled with an acceptance of diversity of approaches (i.e., that they should be used independently); the popular association of mixed methods with triangulation and consequent lack of consideration of integrative strategies; and the view that integration or synthesis of results is an intellectual or ideologically driven activity (which, therefore, occurs independently of data handling). I would argue four further practical reasons why it has not been popular: to achieve integration of data analyses requires a breadth of skills that has not been commonly available in a single researcher, or alternatively a close-knit multi-skilled team; it requires the capacity to imagine and envision what might be possible—to tread new paths—along with the logic (and skills) required to bring that about; students (and others) are frequently encouraged to write results from different components of their studies separately (integration in a dissertation is in the ‘too hard’ basket, or is seen as ‘risky’); and, finally, integration is greatly benefited by data handling technology (computer software) to facilitate the process, which, until relatively recently, has not been readily available. Integrative software is still very much in development, and indeed, software for qualitative analysis, from which much of it is derived, is only now beginning to gain wide acceptance in the academic community.



## Two Major Routes to Integration in Analysis

In asking how does (or might) the use of computer software and processing power facilitate or extend integration of analyses, the key question for this paper relates to this issue of data handling technology. The paper will focus on the more ‘everyday’ possibilities for computer assisted analysis of mixed methods data using spreadsheets or databases, and commonly available qualitative and quantitative analysis software. There is a large and growing range of other analysis techniques and specialist software available to the enthusiastic user, often requiring programming for specific purposes: it is beyond the scope of this paper to review their use here.

I propose that *in terms of data handling*, two major routes to integration underlie the various strategies one might adopt when using software:

1. *Combination* of data types within an analysis, such as when categorical or continuous variables are used both for statistical analysis and as a basis for comparison of coded narrative (qualitative) material. This could occur through using both text and numeric data gathered at the same time, for example through a survey instrument; or using sequentially gathered data, most commonly (as identified by Bryman, 2006) a combination of survey and interview.
2. *Conversion* of data from one type to another for analysis, typically the conversion of qualitative codes to codes used in a statistical analysis, but also, alternatively, through the contribution of quantitative data to a narrative analysis of events, circumstances, or perhaps a life history (Elliott, 2005; Tashakkori & Teddlie, 1998).<sup>3</sup>

Strategies such as data consolidation, blending or merging are likely to involve both conversion and combination.

### Using Software to Combine Numeric and Text Data for Analysis

The first challenge faced by the researcher seeking to combine mixed forms of data and procedures for working with them is one of data management—how to link observational or interview text or open-ended survey responses (i.e., textual data) to demographics, responses to fixed-alternative questions, or other measurements (data in numeric form). Traditionally, brief explanatory comments provided in surveys have simply been ‘eyeballed’ by the researchers looking for illustrative comments; responses to open-ended questions might have been category coded to allow for frequency counts and interrelationship with other

variables; and unstructured text has simply been marked with the demographic characteristics of the interviewee, as additional information to be noted by the researcher working with that text.

The advent of text-handling spreadsheets and databases and, in particular, of text analysis software, has heralded solutions to these data management problems, and opened up new possibilities for more rigorous and/or deeper analysis of this type of data. They have not necessarily solved the theoretical issues which could arise when different forms of data are combined, however.

### *Using a General-Purpose Spreadsheet or Database*

In its most elementary form, integration of data through combination occurs in structured surveys where a pre-categorized (closed) response to a question is followed up with a request to respondents to provide comment, explanation or illustration of their answer. Comments might be sorted by the categorized responses to provide illustrative material to assist in interpreting what each response really meant to the survey respondents. Such sorting is a simple task in any spreadsheet or database, through which all open responses from any given subgroup (demographic, or based on categorical responses to a parallel question) can be brought together and compared with those from a different subgroup. Analysis in such cases rarely extends beyond identification of patterns in the text in relation to respondent groups, although it is also possible to consider patterns of which respondents gave what kinds of answers and to investigate anomalies in the responses, for example, when people who chose contrasting categories of closed response provided the same kind of elaboration of their answers.

Unstructured data can be similarly organized in a spreadsheet by defining a set of issues to explore, and entering brief summaries of what was said by each respondent under each issue (issues in columns, respondents in rows). Data which categorizes respondents are also entered as one or more columns, and are used to sort the textual comments, revealing any patterns in responses which may be present. This is quite a reductionist approach to qualitative analysis (Miles & Huberman, 1994), but is useful where time for analysis is limited or the data lack ‘richness’ and where relevant issues are largely identified before analysis. New categories or issues can be added during the process if found to be necessary, by adding an additional column, or additional categorization of the text summaries can be completed during analysis to allow further sorting and examination of relationships between categories. This method was used with data derived from interviews with heads of academic departments, in six discipline areas across twelve Australian universities, regarding the research career

## THE CONTRIBUTION OF COMPUTER SOFTWARE TO INTEGRATING QUALITATIVE AND QUANTITATIVE DATA AND ANALYSES

opportunities afforded new academic staff in their departments (Bazeley et al., 1996). Sorting of responses revealed that new staff in physics had much greater opportunities given them (“honeymoon periods” from teaching, computer facilities, financial support) and that research activity was “expected,” in comparison with those in nursing where the majority of new academic staff were still undergoing research training, support for research was more patchy and teaching demands were high, while for those in psychology, staff had research qualifications, research was “supported” and necessary equipment was usually available but teaching loads were a problem. Interestingly, patterns were much more clearly defined by discipline than by the status of the university.

### *Using Qualitative Data Analysis Software*

Using qualitative data analysis software (QDAS), when the textual comments warrant more detailed analysis, allows the researcher to take analysis of mixed, structured survey data a step further than is possible using a spreadsheet or database. Assuming appropriate formatting, a number of QDAS programs now have a facility for autocoding text for the question to which it was a response, as well as for importing individual matching statistical data (such as demographics or categories of response to closed questions). This allows the kind of sorting (of text response by value of pre-categorized response variable) any database can do, as outlined above. But, unlike spreadsheets or regular databases, the greater flexibility of coding systems in QDAS means that the text material can also be readily coded into new emergent concepts or categories.<sup>4</sup> Text stored in these new coding categories, also, can be viewed comparatively across demographic subgroups, or in relation to responses to parallel (or other) categorically coded questions. This technique was used to combine analysis of responses to both closed and open-ended questions covering knowledge of and attitude to organ donation, given by those who had been faced with this issue in a personal way (Pearson, Bazeley, Plane, Chapman & Robertson, 1995). Answers to a question on reasons why one might personally choose to donate were coded to create three categories reflecting altruism, pragmatism, and anxiety about the integrity of the body. These then could be considered in relation to grief resolution (and other variables). The patterning of responses was clear for those expressing a pragmatic viewpoint (who were resolved or resigned) or a concern with body integrity (unresolved, or at best, resigned), but those expressing altruism were equally likely to be resolved or unresolved in their grief. Further examination of the sorted text revealed a fresh perspective on the data: all of those unresolved in their grief who expressed altruism did so in life-or-death

terms, for example: “If other people can live, why not?” In contrast, all of those who were resolved in their grief and who expressed altruism did so in quality-of-life terms, for example: “A man would be very selfish if he died with healthy organs and didn't give someone else a chance to lead *a normal life*” (emphasis added).

More generally, using QDAS, the capacity to combine unstructured text (or similar) data with demographic, categorical, or scaled information opens up a range of possible analytic strategies that would be much more difficult to achieve without software. Variable data are combined with coded text by using the values of the variables (which apply to whole cases) to sort the intersecting text for a particular coding category, or a set of categories. This facilitates comparison of how different demographic subgroups might refer to an experience, concept, belief or issue; it allows the researcher to compare experiences or expressed attitudes as they arise in different contexts; it opens the possibility to corroborate or confirm the meaning of scaled scores by matching scale points with text in which participants describe relevant experience. For example, patients recovering from day surgery completed a 10 point visual analogue scale to record the level of pain they were experiencing, and were interviewed also about their experience of surgery and pain (Coll, nd). Their descriptions of their experience of pain could be sorted by the rating they had given for the level of pain experienced. In this way, it could be determined what each point on a pain scale of this type meant for people experiencing it, thus making use of the scale more meaningful for further research.

The interaction of multiple variables in relation to a particular coding category or concept can be achieved through refining the query in a way that is somewhat analogous to use of a two-factor analysis of variance, for example, to examine the interaction of gender and discipline with respect to an element of academic experience. Alternatively multiple interactions can be examined through repeated querying of the data for different subgroups, as in a multi-layered contingency table. The matrix function in NVivo facilitates this kind of comparative querying by allowing multiple comparisons at one time, with or without restrictions on what data are considered within each query, but the end result also can be achieved, albeit a little more tediously, with most QDAS. NVivo was used, for example, to compare expressions of satisfaction (personal pleasure) gained from doing research for male and female social scientists and scientists (Bazeley & Richards, 2000). The sorted text suggested that those in each discipline group gained satisfaction from different sources, while differences were not apparent for gender. Approximately half of the members in each discipline group reported satisfaction

(gaining personal pleasure from engaging in research), but those in the sciences who did so were likely to refer to the sense of agency they experienced in doing research, while most of those in the social sciences made reference to achieving a goal or a task when expressing satisfaction.

#### *Benefits from Combining Numeric and Text Data for Analysis*

Multi-method approaches typically bring quantitative and qualitative sources together by using qualitative comments, interviews, or documentary sources to corroborate, illustrate, or elaborate on the meaning of categorized responses to survey questions and quantified instruments; to provide a basis using one type of data for sampling or instrumentation using the other; or to provoke new thinking. As noted earlier, in most published research this has meant only that the qualitative data are placed alongside the quantitative data for analysis, rather than being integrated with it. Use of a computer program in the process of mixing methods can not only assist in, but greatly extend the use of data gathered for complementary or expansion purposes because such use facilitates matching of different data sources for individual respondents; comments, expressions of attitude, or observations made by a particular person can be matched with their particular rating of their own experience, or their demographic details. The comparison process is therefore refined, providing the basis for comparative pattern analysis, illustrative understanding, and potential also to reveal new (or previously unobserved) dimensions in the data (such as source of satisfaction, in the example above). This strengthening of the comparative process may well be one of the more exciting outcomes of using these techniques for the researchers involved, particularly for those employing grounded theory methodology (Strauss, 1987).

Furthermore, when data are matched in the way described, instances where individuals go against a trend can be readily identified and explored in detail. These cases might be outliers on a statistical measure, deviant cases in qualitative terms, or cases where there is an apparent contradiction in the data from the different sources (Caracelli & Greene, 1993; Miles & Huberman, 1994). For example, from the examination of gender and discipline differences in satisfaction referred to earlier, two social scientists (one male, one female) also expressed agency, while one scientist did not. These cases could be identified, revealing that the two social scientists both worked in experimental psychology (which has more in common, perhaps, with science than social science), and the one scientist's current work was all to do with recording the history and biography of science and scientists (which has more in common with social science than science). It

could be argued, then, that rather than contradicting the observed trend, these apparently discrepant cases added confirmation.

When contradictions or other anomalies arise from an exercise in combining data sources, then like subgroup comparisons, this also has the potential to stimulate analytical thinking beyond simple illustration (serving an initiation purpose for mixing methods). The cause of the contradiction or anomaly might be explained methodologically (an important insight in itself), new substantive understanding could result, or, as with triangulation, it could create the need for further data collection in order to resolve emergent discrepancies (Erzberger & Kelle, 2003; Jick, 1979).

#### Using Software to Convert Coding from Qualitative Data for Statistical Analysis

For as long as any of us can remember, open ended responses to survey questions have been category coded for inclusion in a statistical database (Bazeley, 1999). In my early consulting experience when survey techniques were dominant in social research, I would typically make an initial classification of (several hundred) responses into 40-50 categories, which were then recoded into 6-8 broader categories for analysis. The kinds of issues raised in the examples and responses given would then be related to other quantitative responses in the survey. Recent text-analysis modules for some statistical programs now attempt to automate this process by categorizing the open ended responses based on the co-occurrence of words (e.g., SPSS, Wordstat). Some freedom for manipulation of categories is usually available to the researcher. The categorized responses then can be considered along with other statistical data. Disadvantages in these methods include the 'cost' of coding time for the manual method and the potential for generation of meaningless categories using the automated method. While these processes work satisfactorily for short answer responses which generally deal very briefly with just one or two concerns, they 'fall down' for more complex data. The principal disadvantage in these processes of direct conversion for statistical use, however, is that one loses ready access to the original text as one progresses through the analysis process and, consequently, to nuances in the way people express their concerns.

Relatively recent developments (primarily since 1997) in QDAS have changed this situation somewhat. The frequency with which concepts, categories, or themes have been identified in unstructured data by the researcher-analyst is now readily provided, and a number of programs export individual coding information which, either directly or indirectly, is read as a case by variable matrix in a statistical program,

## THE CONTRIBUTION OF COMPUTER SOFTWARE TO INTEGRATING QUALITATIVE AND QUANTITATIVE DATA AND ANALYSES

hence allowing further statistical analysis. Additionally, in some programs, more complex associations between variables can be exported as a quantified matrix (e.g., as a similarity matrix). The defining characteristic of what is happening, in these instances, is that data are being converted (morphed, transformed) for reporting or for further analysis—a process generally referred to as quantizing the qualitative data (Tashakkori & Teddlie, 1998). Critically, however, ready access to the text which supports the exported numeric information is retained.

### *Counting in Qualitative Analysis*

Counting themes, or instances of a category in a qualitative database, constitutes a very simple form of conversion of data from textual to numeric form. For the majority of studies that develop quantitative reports from qualitative data, the quantitative data generated are just descriptive statistics reporting numbers of themes or categories found (Creswell, 2003; Niglas, 2004). Use of counts communicates more effectively and reliably than does use of vague terms to indicate more or less frequent occurrence of some feature in the text (Miles & Huberman, 1994; Sandelowski 2001). Counts can be seen as reflecting the importance of various emergent themes (Onwuegbuzie & Teddlie, 2003), although it can be argued that frequency and importance are not necessarily synonymous.

Qualitative software programs can readily provide various kinds of counts, including the number of text segments coded at a particular category, the number of cases with coding, or volume measures which might include the total number of characters or words coded, the proportion of text coded, and so on. These might be used as simple counts or proportions and descriptively reported as part of a qualitative write-up. While researchers have often used counts of qualitatively derived themes in their work, measures of volume have typically necessitated having the text broken into predetermined segments for coding to facilitate counting and assessment as a proportion of the total (Chi, 1997). When software is used to facilitate such counting of occurrences, however, it becomes less necessary to break the text into predetermined segments in order to code and count, and the whole measurement process is considerably simplified.

Volume counts (in this case, lines of text) were used, for example, by Holbrook and Bourke (2004) in a study of Ph.D. examiner's reports, to determine the relative emphasis given to major components of the dissertation (e.g., literature, methods, analysis, discussion), as well as the relative amounts that comprised summative versus formative evaluation of the work, as a first step in their analysis of the Ph.D. examination process. This was then followed up with qualitative analyses of the types of comments made

(e.g., Holbrook, Bourke, Lovat, & Dally, 2004). Similarly, a decreasing number of lines of text between occurrences was used by Anderson et al. (2001) to verify the snowballing spread of argument strategies between children working in problem-solving groups.

When subgroups are compared (as described earlier), the resulting analyses provide not only an assessment of the qualitative differences in the coded text between the groups, but also a count of the frequency with which that coded concept was used by members of each group. Each alternative component of the information provided (numbers, text) adds to the analytic picture: how many report and how they report might each be conditioned by (or associated with) the subgroup to which each person (or source) belongs; each type of analysis provides different but complementary information.

### *Converting Qualitative Coding to a Case by Variable Matrix for Statistical Analysis*

When conversion is taken a step further, and codes derived from qualitative data are recorded separately for each case in the data (either as presence/absence of each code or as frequency of occurrence), then one has a case by variable matrix. Such case-coding matrices might be based on the presence or absence of *a priori* categories, or on interpretive coding categories generated during the process of analysis. Assuming satisfaction of necessary statistical assumptions for the processes chosen, this type of matrix provides the basic form of data for most statistical analyses, including hypothesis testing, predictive modeling, and exploratory analyses. It can be used either on its own, or it can be amalgamated with an existing quantitative database for the same cases. Converted qualitative coding was combined with an existing quantitative database in an experimental test of the impact of training through classroom discussions involving collaborative reasoning on children's argumentation (Reznitskaya et al., 2001). Following training, children wrote individual persuasive essays based on a different problem from that discussed in training. The essays were coded for presence of formal argument devices and use of textual evidence. ANOVA and ANCOVA were used to demonstrate that having an argument schema developed through training enabled students to consider and present more arguments, independently of socioeconomic status or vocabulary skills. Detailed text analyses were then conducted on a purposive sample of essays to examine and illustrate argumentation strategies used by the children, revealing that "collaborative reasoning students are generally more successful at generating and articulating an argument, considering alternative perspectives, marshalling text information, and effectively utilizing certain formal argument devices" (p.171).

Conversion of coding for statistical analysis raises a number of issues to be addressed by the researcher: (a) there needs to be sufficient cases (preferably probabilistically rather than purposefully selected) to provide statistically sound samples for the procedures selected; (b) a decision has to be made about whether it is more appropriate to export information reflecting volume of text coded, or simply the presence or absence of a code, and (c) if the qualitative category codes data which are non-directional (e.g., that the issue of the character of a witness was raised, without identifying the conclusion reached), then, depending on the purpose, further coding of the data within that category (to more specific codes, e.g., reflecting a positive or negative assessment) could be necessary before export (Bazeley, 2003).

*Exploratory Statistical Analysis of Patterns of Association in Qualitatively Assigned Codes*

Statistical techniques which include cluster analysis, correspondence analysis, and multidimensional scaling have been fruitfully applied to quantitized qualitative data, to develop or clarify concepts or themes, or to test hypotheses (Ryan & Bernard, 2000, 2003). Sometimes the resulting statistical analyses are, in turn, qualitized as more holistic descriptions are built from the statistical evidence, demonstrating the recursiveness often present in mixed methods analysis. For example, Excel and SPSS were used by Niglas (2004) in a primarily quantitative content analysis of mixed methods studies. She used scales to record variation across 145 mixed methods studies on a range of design characteristics. K-means cluster analysis of the quantitative content analysis variables classified the studies into eight distinctive groups, and the characteristics which best differentiated the groups were calculated. Findings based on the statistical analysis were compared with memo-style notes taken during the initial reading of the studies to generate brief descriptions for each of the eight groups—thus qualitizing the quantitized data which, in turn, had been derived from interpretive (qualitative) reading of text. These eight groups were then used to organize the articles for further statistical analyses and conceptual mapping.

A range of statistical techniques, including several based on patterns of association, are being used in an ongoing concept analysis of research performance (Bazeley, unpublished data). The primary data comprise descriptions given by 295 academics for eight different aspects ('brands') of research performance—descriptions of researchers who are productive, active, recognized, satisfied, approachable, and/or who demonstrate quality, ability, benefit. These have been coded using NVivo to create a set of descriptors. Additionally, basic demographic data are available,

along with each academic respondent's weighting of the importance (or value) of each of these eight aspects of performance for doing research and for assessing research (as interval scales). These additional numeric data have been imported into the NVivo database for use in combination with text responses, and coding based on the descriptions given has been exported from NVivo in a number of forms, each contributing to a different type of analysis. For example:

1. A table showing which respondents used which descriptors overall (a case by variable matrix) when combined with the additional quantitative data is allowing a comparison to determine whether research performance is thought about differently depending on gender, discipline, educational status or level of interest or involvement in research.
2. Descriptors used by each academic respondent for researchers displaying each particular aspect of performance, weighted by the value they assign to that aspect of performance, are being exported in order to contribute to a general model of research performance based on both frequency and weighting of responses. For example, if a description of 'good communicator' is given for a productive researcher, it is likely to be given a higher weighting than if it is given as a description for being approachable, when the total picture provided by the descriptors is being developed.
3. A matrix of the frequencies with which each descriptor was used for each aspect of performance has provided the basis for cluster analysis of performance types, confirming a classic quality-quantity divide in understanding performance, but also revealing that social factors and approachability in particular are seen as being quite outside the general domain of research performance, a conclusion supported also by the importance ratings given to approachability. The form of expression used for each descriptor, according to the type being described, is also being reviewed within the NVivo database. For example, although quality and ability 'hang together' statistically, the text suggested differences in emphasis underlying the way that descriptors, such as having substantive knowledge, displaying originality, or theoretical understanding, are expressed in the context of each of ability or quality (Bazeley, 2001).
4. Multidimensional scaling is being applied to a descriptor-by-descriptor similarity matrix, based on the frequency of co-occurrence of

## THE CONTRIBUTION OF COMPUTER SOFTWARE TO INTEGRATING QUALITATIVE AND QUANTITATIVE DATA AND ANALYSES

descriptors given for each performance type (i.e., the number of times respondents used any pair of descriptors in the same context of describing a researcher of a particular type). This process will identify broader dimensions underlying the concept of research performance held by academics and should lead to a simplified conceptual model of research performance, to feed back into further qualitative and quantitative analyses.

5. Scaled weighting data imported into the NVivo database are being used to compare the form of expression of a description given for each performance type, in relation to the value assigned to that type by the same individual. For example, do respondents use different terms or expressions for, say, methodological understanding in the context of ability, depending on whether they rate ability high or low in importance?

These techniques are all being used in an exploratory way, appropriate to the purpose of exploring and elucidating a concept. Extensions to this work are likely to involve confirmatory strategies.

### *Benefits from Using QDAS in Converting Data for Statistical Analysis*

Integration of analyses using conversion of data is useful in initiating fresh perspectives through exploratory studies, particularly those involving concept analysis; for creation or validation of scaled measures; development of typologies; and for studies attempting to identify predictors of an outcome. Such analyses bring the power of statistical analysis to an inductive project, particularly in exploring the structure of data, while retaining the freedom and power of the qualitative techniques to provide situated meaning. Integration involving conversion of data is useful also in studies designed to test hypotheses (such as those on children's argumentation, described above), or to build predictive models where the foundational data are text (for example, from legal judgments, or case histories).

One of the primary benefits of using qualitative coding as the basis for statistical analysis is that the researcher does not have to pre-determine the categories which will be used for analysis. At the same time, there is no guarantee that all participants in the research process will be equally comprehensive in their discussion of the topic, raising the issue, for example, of whether absence of mention of a topic represents lack of importance, deliberate omission, or a temporary lapse in attention.

Richards (2005) drew a distinction between quantitative coding as data reduction, and qualitative coding as data retention, in particular, the retention of the links between ideas and the data that generate those

ideas. The reduction of text to numbers, as in quantitative content analysis, carries the associated problem "that researchers cannot be sure that the meanings they attach to words on a survey and to the resulting statistical summaries are similar to those held by the respondents; the data have become decontextualized" (Rossman & Wilson, 1994, p. 321). In contrast, the use of QDAS in the generation of codes for statistical analysis carries with it the key advantage that text associated with the codes used is retained in a readily accessible way, thus assisting interpretation of patterns during the process of analysis, validation of conclusions through checking findings back against the qualitative data, and initiation of further qualitative analyses or re-analyses.

### Blending Analytic Strategies: Combination and Conversion Working Together

Integration of data and analyses through an amalgamation of both combination and conversion may be necessary to reconcile "divergent findings, paradox, and contradiction" that can result from mixed methods studies, or indeed, to initiate creative insights through resolution of "dissonance, doubt, and ambiguity" (Rossman & Wilson, 1994, p. 323). Iterative use of alternate analytic strategies and the programs which support them within a single analysis is one form of this type of integration of data. Blending or merging of diverse data sources to create new composite variables which are then fed back into the analysis is another.

This latter strategy was used by Kemp (1999) in her study of the community service needs of spinal injured people. She found dissonance between quantitative data indicating that there was a desperate shortage of community service provision, and qualitative data that suggested ambivalence in the spinal injured population about whether they would access services they had most complained about not having, should they become available. Qualitative coding regarding attitude to use of services was combined with a quantitative variable reflecting current use of services to create a new composite variable. Further quantitative analyses using this variable pointed to a perception of arbitrariness in distribution of community services for the spinal injured population. The computed variable, imported back into the qualitative database, was then used in association with both service satisfaction scales and respondents' qualitative responses about the beneficial and detrimental effects of services to reveal that the quantitative arbitrariness of service provision was, in fact, not so arbitrary, but rather, that services were allocated on the proviso that persons with spinal injuries adopt life plans which met the expectations of

service providers (i.e., to be different rather than ordinary).

In reflecting on this experience, Kemp (2001) saw this process of integration as paralleling the iterative process of protein transfer between the sense and anti-sense strands comprising the double helix of DNA. The image of the unwinding and rebuilding of DNA molecules evokes dissonance and ambiguity, and a transformative, interpretive method that can juxtapose numbers and words to achieve a cohesive, integrated explanation. The use of QDAS in association with statistical software facilitates such juxtaposition of numbers and words to create new variables and new understanding.

### Conclusion

Published reports of studies that truly integrate qualitative and quantitative data sources in analysis are rare, as are those which apply both textual and statistical interpretive techniques to a single data source. Studies that use computers to do so are even rarer.

In this paper, I have not attempted to survey the whole field of integration of data and/or analyses in mixed methods research, nor the full range of computer-based strategies available for such integration. Rather I have concentrated on explaining and illustrating the use of 'off-the-shelf' computer software to achieve a combination of qualitative and quantitative data or analyses, or conversion from qualitative to quantitative coding and analysis, as common strategies for integration. To date, it is developments in software programs for analysis of qualitative data that have contributed most noticeably to researchers' capacity for integrating methods in the ways described in this paper. Indeed, Lyn Richards (2002) has argued that the most radical *methodological* changes that came about with qualitative computing were not in what the computer could do (such as coding), so much as the uses to which it could be put in driving a complex and iterative data interrogation process. Just some of what is currently possible, and the rewards from learning to use software tools, have been illustrated above. Tools are still being developed, a process which is both responsive to and which can lead to new techniques in data analysis. The future is open to imagination, and need.

The lead editors for this article were R. Burke Johnson and Anthony J. Onwuegbuzie.

### References

- Anderson, R. C., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S., Reznitskaya, A., Tillmans, M., & Gilbert, L. (2001). The snowball phenomenon: Spread of ways of talking and ways of thinking across groups of children. *Cognition and Instruction, 19*(1), 1-46.
- Bazeley, P. (1999) Tools and techniques for handling qualitative survey data. *Proceedings of the 3rd International Conference of the Association for Survey Computing, 277-284.*
- Bazeley, P. (2001). Integrated analysis. In J. Blasius, E. de Leeuw, J. Hox, & P. Schmidt (Eds.), *Social science methodology in the new millennium*. Opladen, Germany: Leske + Budrich (CD)
- Bazeley, P. (2003). Computerized data analysis for mixed methods research. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 385-422). Thousand Oaks, CA: Sage.
- Bazeley, P., Kemp, L., Stevens, K., Asmar, C., Grbich, C., Marsh, H., & Bhathal, R. (1996). *Waiting in the wings: A study of early career academic researchers in Australia*. National Board of Employment Education and Training, Commissioned Report No. 50. Canberra: Australian Government Publishing Service.
- Bazeley, P. & Richards, L. (2000). *The NVivo qualitative project book*. London: Sage.
- Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research, 6*(1), 97-113.
- Caracelli, V., & Greene, J. (1993). Data analysis strategies for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 15*, 195-207.
- Chatterji, M. (2004). Evidence on "what works": An argument for extended-term mixed-method (ETMM) evaluation designs. *Educational Researcher, 33*(9), 3-13.
- Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences, 6*, 271-315.
- Coll, A. M. *Experience of pain after day surgery*. Unpublished data, School of Care Sciences, University of Glamorgan, Wales.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative and mixed methods approaches* (Second ed.). CA: Thousand Oaks: Sage.
- Elliott, J. (2005). *Using narrative in social research: Qualitative and quantitative approaches*. London: Sage.
- Erzberger, C., & Kelle, U. (2003). Making inferences in mixed methods: The rules of integration. In

THE CONTRIBUTION OF COMPUTER SOFTWARE TO INTEGRATING QUALITATIVE AND QUANTITATIVE DATA AND ANALYSES

- A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 457-488). Thousand Oaks, CA: Sage.
- 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.
- Holbrook, A., & Bourke, S. (2004). An investigation of PhD examination outcome in Australia using a mixed method approach. *Australian Journal of Educational & Developmental Psychology, 4*, 153-169.
- Holbrook, A., Bourke, S., Lovat, T., & Dally, K. (2004). Qualities and characteristics in the written reports of doctoral thesis examiners. *Australian Journal of Educational and Developmental Psychology, 4*, 126-145.
- Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly, 24*, 602-611.
- Johnson, R. B., & Onweugbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14-26.
- Kemp, L. A. (1999). *Charting a parallel course: Meeting the community service needs of persons with spinal injuries*. Unpublished doctoral thesis, University of Western Sydney, Australia.
- Kemp, L. A. (2001, May). *The DNA of integrated methods*. Paper presented at the annual conference of the Australian Association for Social Research, Wollongong, NSW.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage.
- Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative Health Research, 8*, 362-376.
- Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 189-208). Thousand Oaks, CA: Sage.
- Niglas, K. (2004). *The combined use of qualitative and quantitative methods in educational research*. Tallinn, Estonia: Tallinn Pedagogical University.
- Onweugbuzie, 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. 351-384). Thousand Oaks, CA: Sage.
- Pearson, I.Y., Bazeley, P., Plane, T., Chapman, J.R., & Robertson, P. (1995). A survey of families of brain dead patients: Their experiences, attitudes to organ donation and transplantation. *Journal of Anaesthesia and Intensive Care, 23*, 88-95.
- Rallis, S. F., & Rossman, G. B. (2003). Mixed methods in evaluation contexts: A pragmatic framework. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 491-512). Thousand Oaks, CA: Sage.
- Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2&3), 155-175.
- Richards, L. (2002). Qualitative computing—a methods revolution? *International Journal of Social Research Methodology, 5*, 263-276.
- Richards, L. (2005). *Handling qualitative data*. London: Sage.
- Rossman, G. B., & Wilson, B. L. (1985). Numbers and words: Combining quantitative and qualitative methods in a single large-scale evaluation study. *Evaluation Review, 9*, 627-643.
- Rossman, G. B., & Wilson, B. L. (1994). Numbers and words revisited: Being "shamelessly eclectic". *Quality and Quantity, 28*, 315-327.
- Ryan, G., & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods, 15*(1), 85-109.
- Ryan, G. W., & Bernard, H. R. (2000). Data management and analysis methods. In N. Denzin, & Y. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 769-802). Thousand Oaks: Sage.
- Sale, J. E. M., Lohfield, L., & Brazil, K. (2002). Revisiting the quantitative-qualitative debate: Implications for mixed-methods research. *Quality and Quantity, 36*, 43-53.
- Sandelowski, M. (2001). Real qualitative researchers don't count: The use of numbers in qualitative research. *Research in Nursing and Health, 24*, 230-240.
- Strauss, A. (1987). *Qualitative analysis for social scientists*. Cambridge, UK: Cambridge University Press.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.



Weiss, C. H. (1972). *Evaluation research*. Englewood Cliffs, NJ: Prentice Hall.

Notes

<sup>1</sup> I prefer to use the term *multimethod* to refer to studies in which two or more methods (of any type) are being used such that each retains its distinctive quality, and *mixed method* to refer to studies where the activities associated with each of two or more methods are intertwined or blended prior to final interpretation. In line with common practice, however, I sometimes use the term *mixed method* also in a more generic sense to refer to the general class of studies in which methods are combined in some way or another.

<sup>2</sup> Presented at Sixth International Conference on Logic and Methodology, Amsterdam, August, 2004.

<sup>3</sup> Because the computer plays a lesser role in this type of conversion, and with space limitations, it will not be a focus of discussion in this paper.

<sup>4</sup> Programs differ in whether such coding has to be done directly on the data sources in their original imported form (usually a document for each person), or whether already coded material (e.g., sorted by question asked) can be coded on to new categories. Where both options are available (as in NVivo), choice depends on whether it is more useful to understand all of a person's responses when coding a particular comment, or whether it is more helpful to simply focus on the issue being investigated in that question.

## A Theory-Driven Evaluation Perspective on Mixed Methods Research

Huey T. Chen

*University of Alabama at Birmingham*

*The purpose of this article is to discuss the conceptual framework and strategies used in theory-driven evaluations in relation to mixed methods research and to explore the opportunities and challenges emerging from theory-driven applications. Theory-driven evaluations have frequently applied mixed methods in the past, and these experiences provide some insightful information for future development of mixed methods. In theory-driven evaluations, the application of mixed methods is justified and applied under a conceptual framework called program theory. The conceptual framework of program theory provides a plan and agenda for mixed methods to work collaboratively and de-emphasizes their differences and incompatibilities. Based upon the conceptual framework of program theory, this article provides several strategies for combining qualitative and quantitative methods in theory-driven evaluations. Procedures in applying these strategies are systematically illustrated. Finally, this article discusses challenging issues related to the future development of mixed methods, such as implications of the use of pure versus modified forms of mixed methods and the advocacy of mixed methods research as a "method" paradigm versus a "method use" paradigm.*

Mixed methods research is the systematic combination of qualitative and quantitative methods in research or evaluation. There has been a growing interest in this topic (Johnson & Onwuegbuzie, 2004). Advocates have argued that mixed methods can overcome weaknesses of a single (qualitative or quantitative) method (Greene & Caracelli, 1997; Howe, 1988; Johnson & Onwuegbuzie, 2004; Sechrest & Sidana, 1995). Greene and Caracelli (1997) provided the following major justifications for mixed methods: (a) triangulation: combining qualitative and quantitative methods to study the same phenomenon in order to gain convergence and increase validity (Denzin, 1970), (b) compensatory: using strengths of each method to overcome the weaknesses of the other to enrich the study of a phenomenon, and (c) expansion: using each method to obtain a fuller picture of a phenomenon.

Quantitative and qualitative purists, however, view these two approaches as being based upon incompatible premises and techniques, and argue that mixing methods is neither meaningful nor valuable to pursue (Guba, 1990). Johnson and Onwuegbuzie (2004) have argued that there are some commonalities between quantitative and qualitative methods, and mixed methods research can narrow the divide between quantitative and qualitative researchers, enhancing the quality of a study.

So far, many discussions or debates about mixed methods have been concentrated on philosophical or

methodological issues. The discussion or development of mixed methods also can benefit from experiences based on the application of mixed methods in the field. Practical feedback can provide insightful information about strategies used in combining different methods, and the opportunities and challenges faced in such applications. This type of information could energize the future development of mixed methods. Theory-driven evaluations have frequently applied mixed methods in the past (Chen, 1990, 1997, 2005). The purpose of this article is to discuss some practical experiences of using mixed methods in theory-driven evaluations. More specifically, in this article, I will discuss the conceptual framework and strategies used in theory-driven evaluation that apply mixed methods and the opportunities and challenges emerging from such applications.

### Application of Mixed Methods in Theory-Driven Evaluation

Quantitative and qualitative methods are based on contrasting assumptions and ideologies about social phenomena and knowledge. In order to avoid being accused of proposing a shotgun marriage, mixed methods research must address not only why mixed methods are needed in a study, but also address how these two approaches or methods can be meaningfully combined in a study while minimizing tensions and conflicts. In theory-driven evaluations, the application of mixed methods is justified and applied under a conceptual framework called program theory. The conceptual framework of program theory (Chen, 1990, 2005) provides a plan and agenda for mixed methods to work collaboratively and de-emphasizes their differences and incompatibilities.

---

Correspondence for this article should be addressed to Huey T. Chen, University of Alabama at Birmingham, Ryals Public Health Bldg., RPHB 227K, 1530 3<sup>rd</sup> Ave S., Birmingham, AL 35294-0022. Email: hchen@uab.edu

---

An analogy to the use of mixed methods in theory-driven evaluation is that of two experts, while equally skilled in their crafts, having very different skills and incompatible views and preferences. Because they have incompatible priorities and standards, it is difficult for them to appreciate the other person's skills and potential contributions, making it challenging to work together to accomplish a task. Working together would be possible, however, if there were a master plan that carefully specified their roles and responsibilities for accomplishing the task. The division of labor provided by the plan would allow two experts to contribute their skills effectively and without changing their own world-views or preferences. The conceptual framework of program theory offers such a plan for theory-driven evaluations to use mixed methods. The conceptual framework of program theory serves as a superordinate goal for quantitative and qualitative methods to pursue jointly (Chen, 1997a). The superordinate goal provides opportunities for qualitative and quantitative methods to use their strengths to make contributions to achieving the common goal and de-emphasizes their differences and incompatibilities. By focusing on the superordinate goal, potential conflicts and tensions of qualitative and quantitative methods are minimized.

#### The Conceptual Framework of Program Theory

When key stakeholders design or implement an intervention program, they usually have some ideas about how the program should be constructed and why the program is supposed to work. Program theory is defined as a set of explicit and/or implicit assumptions held by stakeholders about what actions are required to solve a social problem and why the problem will respond to these actions (Chen, 2005). A program theory is the stakeholders' theory. However, stakeholders usually do not clearly and systematically document their program theories. In conducting theory-driven evaluations, evaluators need to facilitate stakeholders' clarification of their program theories. Chen (2005) provides a conceptual framework for program theory that is useful in guiding evaluators in facilitating stakeholders' clarification of their program theories. This conceptual framework is illustrated in Figure 1.

Figure 1 indicates that a program theory consists of two models: an action model and a change model. The change model at the bottom of Figure 1 depicts the causal process generated by the program. The change model consists of the following three components: (a) intervention, which refers to a set of program activities that focus on changing the determinants and outcomes; (b) determinants, which refers to leverages or mechanisms that mediate between the intervention and outcomes; and (c) outcomes, which refers to the

anticipated effects of the program. The change model assumes that the implementation of the intervention will affect the determinants, which in turn, will change the outcomes. Solid arrows in the change model represent casual relationships among the components.

The action model on the top of Figure 1 represents a systematic plan for arranging staff, resources, setting, and support organizations in order to reach target populations and provide intervention services. The action model consists of six components: (a) implementing organizations, (b) program implementers, (c) associate organizations and community partners, (d) ecological context, (e) intervention and service delivery protocols, and (f) target population.

*Implementing organization.* The implementing organization is responsible for organizing staff, allocating resources, and coordinating activities to implement a program. The capability of the organization affects the quality of implementation. Evaluators can assess the capacity of the implementing organization and/or provide information to assist stakeholders in enhancing and ensuring the capacity of the implementing organization.

*Program implementers.* Program implementers are the people who are responsible for delivering services to clients such as case managers, outreach workers, school teachers, health counselors, and social workers. Evaluators can provide useful information to stakeholders in assessing the recruitment and training process and determining implementers' competencies and commitment.

*Associate organizations/community partners.* Programs often benefit from or require cooperation or collaboration between their implementing organizations and associate organizations and community partners. Evaluators can provide useful information to stakeholders by assessing whether or not the program has established the needed collaborations.

*Ecological context.* The ecological context is the portion of the environment that directly interacts with the program. Programs may need support from environments such as social supports and social norms to facilitate program success. Evaluators can provide useful information by assessing whether the program has the needed support from the environment.

*Intervention and service delivery protocols.* An intervention protocol is a curriculum or prospectus that states the exact nature, content, and activities of the intervention. The service delivery protocol refers to the particular steps to be taken to deliver the intervention in the field.

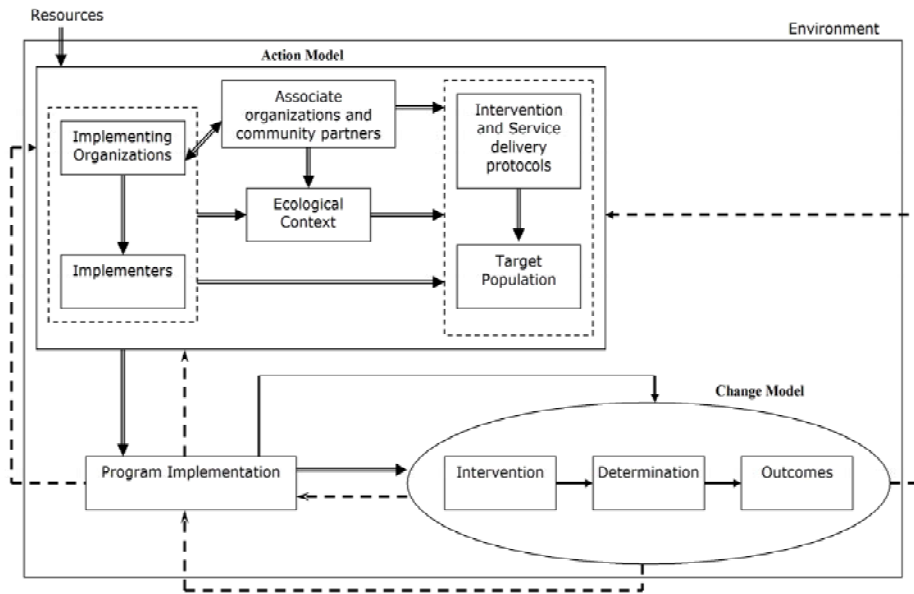


Figure 1. Conceptual Framework of Program Theory (Comprehensive Form)

*Target population.* The target population is the group of people that the program is intended to serve. The success of a program is affected by the following factors: the presence of well-established eligibility criteria, the feasibility of reaching eligible people and effectively serving them, and the willingness of potential clients to become committed to or cooperative with the program. Evaluators can provide useful information for assessing the adequacy of a program in identifying, screening, and serving target populations.

The double-banded arrows between two components within the action model represent the sequential order between the components. That is, the completion of some component provides the basis for completing the next component. For example, the double banded arrow from “implementing organization” to “implementers” indicates that it is usually a requirement to have a capable implementing organization in place in order to recruit and train implementers adequately.

The action model and change model are closely related to each other and are essential for the success of a program. On the one hand, a change model is needed to justify the selection of an intervention for achieving the goals and/or outcomes and it provides a basis for developing the action model. On the other hand, the action model provides a blueprint to organize program activities and to activate and energize the change model for achieving program goals.

Guided by the conceptual framework of program theory, theory-driven evaluations provide a holistic assessment of a program. Unlike a black-box evaluation that provides, mainly, information about the relationship between an intervention and outcomes, a theory-driven evaluation provides rich information about how and why a program reaches or fails to reach its program goals. The comprehensive information provided by theory-driven evaluations also is useful to stakeholders in meeting their program accountability and improvement needs.

#### Theory-Driven Evaluations Strategies for Applying Mixed Methods

In theory-driven evaluation, the need for program theory clarification and holistic assessment usually requires the use of mixed methods. Theory-driven evaluations require two primary tasks: (a) facilitating stakeholders in clarifying or developing their program theory, and (b) empirically assessing program theory. The comprehensive scope of theory-driven evaluation involves the sequential combination of these two tasks, and the program theory shows the needs and opportunities for using various strategies for applying mixed methods. Table 1 lists several strategies for using mixed methods in theory-driven evaluations. According to Table 1, program theory clarification can be conducted by either qualitative or quantitative

methods. Empirical assessment can be qualitative, quantitative, or mixed. Different combinations of methods for program theory clarification and for empirical assessment require different strategies. Because qualitative methods have been the popular mode for clarifying stakeholders' views of theory, this paper mainly focuses on the top half of Table 1.

*Switch Strategy*

In this switch strategy, one first applies qualitative methods to clarify stakeholders' program theory and then uses quantitative methods to assess the program theory. The switch strategy (qualitative then quantitative sequence) also is frequently used in theory-driven outcome evaluation to assess the change model (Chen, 1990). The procedures related to such an application are: (a) intensive interviews, and (b) working groups.

*Intensive interviews.* In this research mode, an evaluator carries out one-to-one interviews with key stakeholders to facilitate making explicit their assumptions about the major program components, processes, and outcomes. This assists them in explicitly formulating their program theories. This method is often used with small-scale programs.

*Working groups.* Large-scale programs tend to have many diverse and often vocal stakeholders. In such programs, program theory needs to be developed

in an open and inclusive manner in order to achieve buy-in from the various stakeholder groups. In this case, the working group method is the preferred mode for evaluators to facilitate stakeholders in developing the program theory. A working group consists of representatives from different stakeholder groups. Group members need to include those who are deeply involved in designing the program and other key constituencies whose input will be influential as to the direction the program will take. Evaluators serve the role of facilitator in the meeting.

After the change model has been established, quantitative methods such as experimental and quasi-experimental designs (Shadish, Cook, & Campbell, 2002) and statistical methods such as path analysis or structural equation models (Maruyama, 1998) are often used to assess the model. For example, a comic book containing anti-smoking messages was used as an intervention in a school-based anti-smoking program. Evaluators used intensive interviewing to assist key program stakeholders in clarifying their views of the change model (Chen, Quane, Garland, & Mracin, 1988). According to key program stakeholders, the reason that the comic book was supposed to have an impact on reducing pro-smoking attitudes and behaviors was because the youngsters would keep and read the comic book. Youngsters would be familiar with the characters, including the anti-smoking hero,

Table 1  
Strategies for Applying Mixed Methods in Theory-Driven Evaluation

Program Theory Clarification	Empirical Assessment	Strategies
Qualitative	Quantitative	Switch
	Mixed methods for different elements or components	Complementarity
	Mixed methods for including contextual information	Contextual overlaying
	Mixed methods for triangulation	Triangulation
Quantitative	Qualitative	Switch
	Mixed methods for different elements or components	Complementarity
	Mixed methods for including contextual information	Contextual overlaying
	Mixed methods for triangulation	Triangulation

the heroine, and tobacco villains, and they would remember the story about warring between these two camps. The information was expected to increase youngsters' anti-smoking attitudes and reduce their smoking. The change model was assessed by using the non-equivalent comparison-group design and path analysis (Chen et al., 1988).

The switch strategy also can be applied to process evaluation. For example, the action model for a school violence prevention program could be clarified using qualitative methods, and the model might end up including components such as teaching students using a prevention curriculum, student group projects on violence prevention, and parental involvement in prevention activities. Based on the action model, evaluators could use quantitative methods. For example, a survey could be used to ask teachers about their implementation of curricula, students about their participation in group projects, and parents about their participation in various school activities.

In the switch strategy, qualitative and quantitative methods are carried out at different times and in a sequential order. In such an application of mixed methods, there is no requirement to alter or change the components and procedures of either the qualitative methods or the quantitative methods. Qualitative and quantitative methods are applied as intact approaches.

#### *Complementarity Strategy*

As shown in Figure 1, the conceptual framework of a program theory consists of multiple components. Some elements within a particular component might require quantitative information and other elements might require qualitative information. The use of quantitative and qualitative methods also varies across the different components. The complementarity strategy is the use of qualitative and quantitative methods to collect different pieces of information for assessing a program theory in order to gain clear understanding of a program. For example, the implementer component might require a quantitative assessment of implementers' competencies, whereas the ecological context component might require having face-to-face interviews with key informants in a community to assess community support for the program. In addition, some elements within a single component might require quantitative assessments, while other elements require qualitative assessments. For example, in the target population component, quantitative data are needed to assess the number of required sessions participated in by the clients, whereas, qualitative methods such as focus group meetings are used to gain an understanding of their views and satisfaction with the program. In the complementarity strategy, a division of labor is used for quantitative and qualitative methods to work and

provide a fuller picture and a better understanding of a program.

Theory-driven process evaluation and integrative process/outcome evaluation are two types of theory-driven evaluations that frequently apply the complementarity strategy.

*Theory-driven process evaluation.* Theory-driven process evaluation is the comprehensive assessment of the implementation of the action model. As indicated in Figure 1, the action model consists of major components such as the intervention and service delivery protocols, implementing organization, implementers, target population, associated organizations, ecological context, implementation process, and the interrelationships among these components. Depending on the program, some components might need to be assessed by using qualitative methods while others might need the use of quantitative methods. Theory-driven evaluations use qualitative and quantitative methods to collect information on different elements of a program in order to obtain a full understanding of a program.

For example, in a comprehensive evaluation of a workplace smoking policy, Gottlieb, Lovata, Weinstein, Gree, and Eriksen (1992) used quantitative methods, including administration of a survey, to collect social and demographic information about the employees and their overall opinions on policy development, implementation process, compliance, and policy impacts. Qualitative methods such as focus groups, individual interviews, and document review were used to gather information about whether the policy had affected social relationships between smokers and nonsmokers and how infractions had been managed.

Mixed methods also were used in a comprehensive evaluation of the implementation of a school-based anti-drug abuse intervention program (Chen, 1997b). One of the components in the program was a requirement for middle school teachers to serve in a counseling role. In assessing that component, quantitative methods were used to rate teachers' satisfaction with a workshop on drug counseling skills sponsored by the education ministry, whereas qualitative methods were used to probe contextual issues of teachers' experiences with the training and service delivery. The qualitative data indicated that teachers felt they were already overloaded with teaching duties and were concerned about the requirement that they take on the additional counseling roles. These qualitative data helped to explain the quantitative findings which indicated that the teachers were not enthusiastic about the training.

*Integrative process/outcome evaluation.* An integrative process/outcome evaluation is a systematic combination of process and outcome evaluation in a

single evaluation. Due to the comprehensiveness of this type of evaluation, it usually requires application of both qualitative and quantitative approaches to evaluate the action model and the change model as shown in Figure 1. For example, the evaluation of a garbage reduction program in Taiwan (Chen et al., 1997) was an integrative process/outcome evaluation. Garbage was collected by government sanitation workers on a daily basis in Taiwan. In order to reduce the ever-increasing amount of garbage, a policy of no garbage dumping and no garbage collection on each Tuesday was established for a community to encourage residents' garbage reduction. In the process evaluation, evaluators used field observations to monitor and determine whether the residents were adequately informed about the new policy and whether the sanitation workers were diligently preventing residents from dumping garbage on Tuesdays. Quantitative methods were used to assess the number of residents who violated the policy and received verbal warnings or fines. In the outcome evaluation, a multiple-group interrupted time-series design (Cook & Campbell, 1979) and a survey were used to assess the change model and determine whether the policy encouraged residents to produce less garbage in the community.

In this study, qualitative and quantitative methods were used to collect information on different elements and components of the action model and quantitative methods were used to assess the change model. The combination of qualitative and quantitative data provided a better understanding of how the program theory worked in the field.

#### *Contextual Overlaying Strategy*

Evaluative data of a program might not be clear or consistent. The contextual overlaying strategy refers to the use of a method (qualitative or quantitative) to collect contextual information for assisting in interpreting the data or reconciling inconsistent findings provided by the other method (qualitative or quantitative). A popular use of this strategy in theory-driven evaluation is to use qualitative methods to collect contextual information for assisting in the interpretation of quantitative data or reconciling findings. Theory-driven evaluation frequently collects qualitative data on the multiple components of an action model. The qualitative data often provide excellent contextual information for assisting in interpreting the quantitative process or outcome data. For example, in an evaluation of a summer job assistance program for urban youths of poor families, the quantitative data from the attendance records (the intervention protocol component) showed that many of youths failed to show up for work or could not hold on to their jobs in spite of their liking the program. The quantitative findings needed additional contextual

information to help one understand their meaning. The qualitative data collected in the ecological component also might help in interpreting the data. For example, a face-to-face interview of the youths might find that many of the jobs were located far away from the youths' homes. Perhaps they had no reliable transportation to get to work on a regular basis.

The contextual overlaying strategy also can be applied in a sequential format. For example, additional qualitative inquiry could be carried out after completion of a quantitative assessment to assist in making interpretations about the quantitative process or outcome findings. For example, a quantitative outcome evaluation indicated that a free fertilizer-to-farmer program in a developing country failed due to low participation of farmers. A field study was carried out to study the reasons for the program failure. It was found that the problem was in the service delivery protocol. The fertilizers were distributed at local police stations. Unfortunately, farmers and other local residents were afraid of and distrusted the police. They were reluctant to go to police stations to provide their names and other personal information to get the free fertilizers.

#### *Triangulation Assessment Strategy*

The triangulation strategy is an application of multiple methods or mixed methods in cross-validating an observed phenomenon. The triangulation strategy can be viewed as a sort of switch strategy except that after the development of the program theory, mixed methods are used to assess the program theory. The triangulation strategy can enhance the validity of an assessment. There are two kinds of triangulation in theory-driven evaluation. The first kind is the multiple qualitative or quantitative method triangulation or simply "multiple methods" triangulation. The second kind is the mixed, qualitative and quantitative, methods triangulation or simply "mixed methods" triangulation. Theory-driven evaluation has used both kinds of triangulation. For example, there has been a great deal of interest in using physicians to deliver HIV prevention messages to HIV-positive patients in clinical settings for the purpose of preventing the spread of HIV. In a process evaluation of the physician-based prevention program (Chen, Grimley, Aban, Waithaka, & Bachmann, 2006), stakeholders wanted to know the fidelity quality of intervention, that is, whether physicians who participated in such a program actually delivered the services (intervention protocol component). This issue was important because HIV prevention was new to the physicians and it was of interest whether they would actually deliver prevention services, even if they agreed to participate in the project. The intervention fidelity was assessed through the use of multiple quantitative methods

triangulation. After each clinical session, the physician was required to file a form documenting whether HIV prevention was delivered according to the protocol and recording the length of time spent discussing HIV prevention with clients. In an exit survey, each patient was interviewed by a research staff member and was asked the same questions. In this way, each patient's report provided a cross-validation of the physician's report on the intervention fidelity. The data from both sources indicated that the services were delivered in the majority of the treatment sessions.

According to stakeholders' program theory, physicians' buy-in to HIV prevention also is an important issue for sustaining the program. The assessment of physicians' buy-in to HIV prevention was based on the use of the mixed methods triangulation strategy. A quantitative assessment of physicians' commitment was carried out by asking physicians to answer a standardized set of survey questions. A qualitative assessment of physicians' commitment included the use of participant observations. Two project coordinators who worked with physicians on a day-to-day basis were asked to keep work diaries documenting physicians' activities and recording their reactions toward the project. These diaries provided rich information on whether the physicians were cooperative or supportive of the intervention in their day-to-day activities. The preliminary qualitative and quantitative data consistently indicated that physicians delivered the prevention services, but their commitment to HIV prevention was not very high during the first six months of the project.

It is interesting to point out that there have been fewer applications of the mixed methods triangulation strategy in outcome evaluation than in process evaluation. One of the major reasons is the costs involved. For example, it is already highly expensive to apply a quantitative method such as a randomized experiment in an outcome evaluation for a program such as providing HIV prevention services to injection drug users. The application of the mixed methods triangulation strategy would mean sending a group of qualitative evaluators into the field for a long time to engage in a prolonged investigation of the effect of the program. The qualitative outcome evaluation could easily be as costly as the quantitative outcome evaluation. The application of the mixed methods triangulation in outcome evaluation could easily double the price of an evaluation in comparison to the use of a single method. This additional cost is a huge barrier for funding agencies and other decision-makers who are deciding if they want to support an outcome evaluation that uses the triangulation strategy. This barrier would be more easily overcome if there were more evidence available to use in convincing funding agencies that the

benefits of such applications offset the costs. One viable alternative would be for mixed methods methodologists to devise innovative mixed methods triangulation strategies for outcome evaluations that are moderately priced. These issues deserve future discussion and investigation.

#### Challenging Issues for Mixed Methods Evaluation and Research

##### *Pure Versus Modified Form Mixed Methods*

When qualitative and quantitative methods are combined into mixed methods research, the two methods can retain their original structures and procedures (pure form) or they might need some adaptations to fit the research and cost situations (modified form). Some implications of pure form versus modified form mixed methods in theory-driven evaluations are discussed next.

*Pure form mixed methods.* Mixed methods under the switch strategy are of a pure form. In this strategy, qualitative methods are used to facilitate stakeholders' clarification of their program theories. On the other hand, quantitative methods are used to assess the program theory. There is no need to modify the procedures of qualitative and quantitative methods. In other words, qualitative and quantitative methods are used in their pure forms under the switch strategy. Furthermore, the expensive part of an evaluation is the assessment side of data collection. Under this strategy, because qualitative methods do not involve collecting data for assessment, but the quantitative methods do, the methods are not both competing for assessment resources.

Strategies such as complementarity, contextual overlaying, and triangulation require the use of mixed methods for data collection. There are pros and cons for use of these strategies in pure form, mixed methods data collection. One major advantage is that the pure form type of application ensures the original integrity of the qualitative and quantitative methods. However, a major disadvantage for such an application is its expense. Qualitative and quantitative methods compete for evaluation resources. Cost is a big hurdle that might prevent widespread use of methods in their pure forms. The most expensive situation is the application of pure form mixed methods under the triangulation strategy to conduct an outcome evaluation. In a sense, the outcome data, which are highly time-consuming to collect, are independently collected twice to investigate the same phenomenon. Furthermore, it is hoped that the results from qualitative and quantitative data will show convergence. Tensions of interpretation arise when the results from the qualitative and quantitative methods are not convergent. In this situation, the evaluator will need to rely on using the contextual



overlaying strategy for collecting contextual information which should help in reconciling the differences or non-convergence.

*Modified form mixed methods.* When the budget and/or timeline prohibit the use of pure form, the modified form of mixed methods is a viable alternative. In the modified form, evaluators alter the structure and procedures of the qualitative and/or quantitative methods or use less expensive or less rigorous applications to save money or meet restrictive timelines. Using the complementarity strategy as an example, the modification could be with respect to the qualitative methods, the quantitative methods, or both. For example, on the qualitative methods side, the application of a case study might be carried out as a site visit for a few weeks rather than a rigorous field study lasting several months or years. On the quantitative side, a survey instrument might be used with a hundred clients rather than a thousand, which might have been needed to obtain a large representative sample. In this situation, the modified forms of qualitative or quantitative methods used in the complementarity strategy are weak if they are judged individually from the perspective of a mono-qualitative or a mono-quantitative methodology. The merits of the modified form mixed methods in the complementarity strategy emerge in their totality. In other words, a new set of criteria needs to be established for judging the quality of the modified form of mixed methods evaluation and research. This new set of criteria has not been systematically discussed or developed yet in the literature. Future studies in this area are greatly needed.

#### *Mixed Methods as a "Method" Paradigm or "Method Use" Paradigm?*

Mixed methods research indeed has merits as has been demonstrated in the applications of this approach in theory-driven evaluation. An important issue is how to position mixed methods in the arena of research methodology. Advocates of mixed methods have different views on this issue. For example, some might argue for mixed methods as an additional method paradigm to the existing qualitative or quantitative paradigm. Others might claim that mixed methods is a new method paradigm that is superior to the qualitative or quantitative paradigm and perhaps can be used to replace those "outdated" approaches.

Based upon my experiences in applying mixed methods in theory-driven evaluations, I propose another view on this issue. I advocate mixed methods as a "method use" paradigm rather than a "method" paradigm at its current stage of development. To qualify to be a method paradigm, a method tradition must have a body of their own unique research methods for researchers to apply. Qualitative or quantitative methods are qualified to be a "method"

paradigm because they have an existing body of their own unique methods. However, currently this is not true for the mixed methods tradition. To date, the mixed methods tradition does not have its own unique set of methods. Instead, mixed methods research relies on combining qualitative and quantitative methods for carrying out research. It would be questionable for mixed methods to claim that it is a "method" paradigm. Instead, perhaps it would be more justifiable and less controversial at this time to call mixed methods a "method use" paradigm to reflect this current situation. The advantages for such advocacy includes: First, it would reduce unnecessary conflicts between mixed methods advocates and qualitative or quantitative methods advocates. Second, it points out a great need for systematically developing mixed method "use" strategies as well as establishing its own standards and criteria for assessing the method use. Third, it highlights the ultimate goal of mixed methods research as being to develop its own unique methods. When mixed methods research has its own body of unique methods, we could then move mixed methods from a "method use" paradigm to a "method" paradigm.

#### Summary

This article describes the conceptual framework of program theory used in theory-driven evaluations employing mixed methods. Under this framework, the following strategies have been used to combine systematically qualitative and quantitative methods in an evaluation: (a) switch, (b) complementarity, (c) contextual overlaying, and (d) triangulation. Benefits and challenges in applying these strategies were systematically illustrated. Two types of mixed methods were also distinguished: pure and adapted. Because the mixed methods tradition still lacks its own body of unique methods that distinguish it from qualitative and quantitative methods, the article advocates for mixed methods as a "method use" paradigm rather than a "method" paradigm at the current time.

The lead editors for this article were R. Burke Johnson and Anthony J. Onwuegbuzie.

#### References

- Chen, H. T. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.
- Chen, H. T. (1997a). Applying mixed methods under the framework of theory-driven evaluations. In J. Greene & V. Caracelli (Eds.), *Advances in mixed methods evaluation: The challenge and benefits of integrating diverse paradigms* (pp. 61-72). San Francisco: Jossey-Bass.

- Chen, H. T. (1997b). Normative evaluation of an anti-drug abuse program. *Evaluation and Program Planning* 12, 195-204.
- Chen, H. T. (2005). *Practical program evaluation: Assessing and improving planning, implementation, and effectiveness*. Thousand Oaks, CA: Sage.
- Chen, H. T., Grimley, D., Aban, C., Waithaka, Y., & Bachmann, L. (2006, in progress). *An evaluation of the implementation of a computer-based, provider-delivered HIV prevention intervention in the primary care setting*. Manuscript in preparation.
- Chen, H.T., Quane, J., Garland, N., & Mracin, P. (1988). Evaluating an anti-smoking program: Diagnostics of underlying causal mechanism. *Evaluation and the Health Professions*, 11, 441-464.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Denzin, N. K. (1970). *The research act in sociology*. London: Butterworth.
- Gottlieb, N. H., Lovata, C. Y., Weinstein, R., Gree, L. W., & Eriksen, M. P. (1992). The implementation of a restrictive worksite smoking policy in a large decentralized organization. *Health Education Quarterly*, 19(1), 77-100.
- Greene, J., & Caracelli, V. (1997). *Advances in mixed methods evaluation: The challenge and benefits of integrating diverse paradigms*. San Francisco: Jossey-Bass.
- Guba, E. G. (1990). The alternative paradigm dialog. In E. G. Guba (Ed.), *The paradigm dialog* (pp. 17-27). Newbury Park, CA: Sage.
- Howe, K. R. (1988). Against the quantitative qualitative incompatibility thesis or dogmas die hard. *Educational Researcher*, 17(8), 10-16.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Maruyama, G. M. (1998). *Basics of structural equation modeling*. Thousand Oaks, CA: Sage.
- Sechrest, L., & Sidani, S. (1995). Quantitative and qualitative methods: Is there an alternative? *Evaluation and Program Planning*, 18, 77-87.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton- Mifflin.

## Enhancing the Policy Process Through the Use of Ethnography and Other Study Frameworks: A Mixed-Method Strategy<sup>1</sup>

Valerie J. Caracelli

*U.S. Government Accountability Office*

*This article examines the current context of federal program evaluation and the prominence given to evaluation for program accountability purposes. Mixed-method studies have been recognized for their capacity to strengthen evaluation evidence. While ethnographic studies can be crafted as mixed-method designs, it is also the case that they can be added as another component or integrated into a different study framework (such as a randomized experiment or survey) creating other mixed-method designs. Several examples illustrate how together the contrasting method frameworks can complement and strengthen the evaluation evidence, increase accuracy, provide important information about context, and enhance explanation and confidence in findings. A pragmatic approach to mixed-method evaluation that weighs the resource issues and potential benefits in using such designs is recommended.*

The ushering in of the 21st century brought unexpected challenges to the United States with 9/11, changing population demographics burdening the health care system, new technologies, program efforts focused on combating terrorism, and issues related to coordination and cooperation among agencies in disaster assistance (e.g., the federal response to Hurricane Katrina). Tensions proliferate in current evaluation practice. For example, today the capacity for electronic record linkage, while offering potential benefits in tracking program outcomes and informing policy debates, can also create new information at the level of the individual that has the potential to violate an individual's right to privacy. Current demands to demonstrate clear evidence of program effectiveness fosters a competitive programmatic environment given diminished evaluation resources (U.S. Government Accountability Office [U.S. GAO], 2001; U.S. GAO, 2005a; U.S. GAO, 2005b; Lane, 2006).

It is in this tumultuous context that evaluation takes place. Evaluation studies are expected to produce information that can be used to make critical decisions about programs so that scarce resources are invested in programs that benefit the country and its citizenry, and problematic programs are improved or replaced. Federal evaluations frequently rely on multiple and mixed-method studies capable of providing defensible evidence coupled with a sophisticated understanding of

the context. Ethnography is particularly suited for providing a framework for examining not only purposeful results-based use, conceptual use, or political or persuasive use of evaluation results but also sources of influence that emanate from the process of conducting the evaluation. (See Kirkhart, 2000, for an explication of an expanded understanding of evaluation use that is modeled as an integrated theory of influence.) Further, process-based influence has a political dimension that can reveal dynamics of power and privilege in the evaluand, provide a forum for all voices and opportunity for constructive dialog, or, at the least, provide the potential for raising issues and drawing attention to social problems (Greene, 1988).

The present paper examines the shift in evaluation toward accountability, the concomitant preference for designs—such as randomized controlled field trials—that have the potential to determine whether desired effects, if obtained, are attributable to the program, and examines the utility of mixed-method studies that include ethnographic methods to assure contextual understanding.

### Shift in Federal Evaluation Toward Evaluation for Accountability Purposes

Evaluations may be conducted for a number of different purposes that are generally characterized as falling within three general perspectives: (a) evaluation for *development* purposes emphasizing building capacity and improving institutional performance; (b) evaluation for *accountability* purposes involving the measurement of results or efficiency, and (c) evaluation for *knowledge* purposes to develop a deeper

---

Correspondence for this article should be addressed to Valerie J. Caracelli, U. S. Government Accountability Office, 2727 29<sup>th</sup> St. N.W Apt. 512, Washington, DC 20008.  
Email: CaracelliV@gao.gov

## ENHANCING THE POLICY PROCESS THROUGH THE USE OF ETHNOGRAPHY AND OTHER STUDY FRAMEWORKS: A MIXED-METHOD STRATEGY

understanding of the factors underlying social problems and the theory and logic inherent in programmatic solutions (Chelimsky, 1997). At the federal level, the prominence given to evaluation for program accountability purposes was already underway in the 1990s when Congress and the executive branch put in place a statutory and management framework for strengthening government performance (U.S. GAO, 1998; U.S. GAO, 2004b). The various reform initiatives such as the Government Performance and Results Act of 1993 (commonly referred to as GPRA or the Results Act; P.L. 103-62) shifted the emphasis from the characteristics of program constituents and the services they receive (process issues) to an increased demand for information on program effectiveness. This shift in perception about what counts as useful information for decision-making was mirrored at state and local levels (U.S. GAO, 1994).

The accountability and program results orientation of federal program evaluation has been reinforced over the last several years with the use of the Office of Management and Budget's (OMB) Program Assessment Rating Tool (PART). PART comprises a standard series of questions used by budget examiners to allow for a consistent approach to assess federal programs in the budget formulation process. PART draws on program performance and evaluation information, scores agencies as to the effectiveness of agency programs and recommends improvements in program design, management, and assessment (U.S. GAO, 2004a; U.S. GAO, 2005c). Further, OMB's guidance to agencies specifies a desire for strong evidence of a program's effectiveness noting this is best achieved through randomized controlled trials (RCTs). Even though the revised guidance acknowledged that RCTs are not always possible or practical to carry out, it reinforced the description of RCTs as the "gold standard" for assessing an intervention's effectiveness in such fields as medicine, welfare and employment, psychology, and education (OMB, 2004).

However, contentious views about the practical capabilities and limitations of RCTs, compared to other designs, have been expressed (Brass, Nunez-Neto, & Williams, 2006). Without clear guidance on alternative methods and the situations for which they are suited, the perception that RCTs were preferred and that strong evidence is necessarily quantifiable evidence was manifested in agencies' responses to accountability for results. In fact, a vitriolic debate among experts within the American Evaluation Association (AEA) took place after the U.S. Department of Education's Institute of Education Sciences took a stance on prioritizing funding for

experimental and some types of quasi-experimental designs over other methods. (For a recap of the debate on whether RCTs should be privileged in funding competitions see Donaldson & Christie, 2004.) In another voice, Chatterji (2004) advocated for the use of multiple methods, effectively implemented, in order to secure quality evidence from studies providing causal explanations of "what works" in education.

The U.S. Government Accountability Office is a legislative branch agency that conducts evaluations, among other types of studies, to improve federal programs and to assure their accountability to Congress and the U.S. Public. A recent study was initiated to determine agencies' progress in implementing OMB's PART-related recommendations for evaluation. The study reported that some program managers disagreed with OMB on the scope, purpose, and quality of their evaluations as well as the usefulness of evaluations by independent third parties unfamiliar with their programs. Several program officials noted that, in designing their evaluations, they were more concerned with learning how to improve their programs than in meeting an OMB checklist (U.S. GAO, 2005d). Federal Evaluators, an informal association of evaluation officials in the U.S. Government, believed that OMB guidance materials reflected too narrow a range of rigorous evaluation designs. Federal Evaluators shared presentations of alternative designs with OMB staff better suited to various program purposes and situations where randomized field trials may not be appropriate (Bernholz et al., 2005).

### Holding Evaluation Designs Accountable

At the federal level, randomized field trials have a long history of use in assessing program effectiveness (Boruch, 1997, 2005); equally illustrious is the history of the use of ethnography in government studies. Both types of designs have been criticized because they are resource intensive, yet both have also been used productively together as a particular type of mixed-method design. The randomized experiment is notably strong in internal validity but constrained by experimental conditions that may limit generalization; moreover, such studies are not always immune to sources of bias such as differential reactivity and biasing social interactions (see Droitcour & Kovar, in press; Fetterman, 1982). The randomized experiment can be strengthened through cross-design techniques that, for example, may use other sources of data to adjust for bias (Droitcour, 1997). Adding or embedding an ethnographic component can provide a more comprehensive study that uses the ethnography to respond to particular study questions, to detail

unfolding explanatory processes during program implementation, or to offer a means by which a more contextualized understanding and thorough explanation of experimental findings can be achieved (Cook, 2001; Cronbach, 1982; see also Onwuegbuzie & Johnson, this volume, for a range of quality issues specific to mixed-method evaluation research).

#### Crafting Mixed-Method Evaluations Using Ethnography and Other Study Frameworks

Mixed-method designs are now recognized for their capacity to serve a number of purposes that strengthen evaluation evidence. Many examples of such designs have appeared in the literature, but there is no hierarchy of preferred designs yielding the strongest evidence. Nor would this be a reasonable expectation. It is precisely the flexibility to craft the best design options, including both qualitative and quantitative forms of evidence, for specific problems and questions that is the strength of mixed methods designs. Such designs may be intra-disciplinary or inter-disciplinary. As Morse (2003) pointed out, an ethnography with methodological congruence can include both qualitative forms of evidence (e.g., in the form of interviews, observations, diaries) and quantitative evidence (e.g., psychometric tests and scales, biological measurements) to answer a research question or enrich explanation. Studies rooted in other social science disciplines that call on methods not characterized as ethnographic (e.g., studies in health, child development, education, criminal justice) can add ethnography to their methods-base to allow for a more in-depth understanding of the phenomena under study. So, although an ethnographic study can be crafted as mixed-method in its design and execution, such a study can also be added as another component or integrated into a different study framework (e.g., randomized experiment, survey).

In previous mixed-method work, the benefits of ethnographic case studies as a complement to quantitative approaches has been illustrated. Ethnographic methods were used to deeply probe quantitative findings and increase understanding and confidence in them (Caracelli & Greene, 1997; Greene, Caracelli, & Graham, 1989). Examples included examining extreme cases that were revealed through regression analysis (Fry, Chantavanich, & Chantavanich, 1981), embedding ethnography into a study using a path-analytic framework in order to examine the influences of culture and environment (Jacob, 1982), and nesting one design within another to illuminate causal processes and increase the interpretability of experimental outcomes (Maxwell, Bashook, & Sandlow, 1986).

For example, in the context of cross-cultural educational research in Thailand, schools depicted as outliers when educational outcomes were analyzed in a multiple regression analysis were further studied using ethnographic case studies for the purpose of clarifying the initial theoretical model (Fry et al., 1981). Ethnographers blind to the quantitative findings were sent to “extreme case” schools to study their educational processes. Ethnographic variables derived from insights not considered in the initial analysis, such as differences in teaching methods, principal characteristics, community support, or other factors important in assessing school effectiveness were incorporated back into the regression model to increase its explanatory power. The technique of “ethnographic residual analysis” allowed for a better understanding of the complex network of relationships among educational inputs, processes, and outcomes (Fry et al., 1981). The potential of such a mixed-method study and integrated analysis strategy is in its capacity to refine, recast, or elaborate initial theoretical assumptions (Caracelli & Greene, 1993).

#### Providing Policy Relevant Information Using Multiple Methods and Ethnographic Case Studies

Today, with an emphasis on program performance and results, it has become incumbent upon agencies to provide evaluation results that demonstrate how well programs are working and/or how to improve agency programs, policies, and procedures. Datta (1997) illustrated how the GAO evaluation of the H-2A Farmworkers program required a mixed-method approach with an ethnographic component in order to be fully responsive to questions needed for a hearing on changes in immigration law. The H-2A program permitted U.S. growers of perishable crops to hire temporary farmworkers from other countries so long as no U.S. citizens or legal immigrants were available to pick the crops. Among other questions, the evaluation focused on whether the H-2A program successfully protected U.S. citizens and legal immigrant farmworkers’ employment opportunities, wages, and working conditions. A multidisciplinary team and several bilingual staff were assembled to implement a multiple methods study consisting of ethnographic case study, historical analysis, field test of worker availability, technical reviews, and secondary analysis of wage data. (The complete methodology can be found in U.S. GAO 1998b.)

The evaluators found only surface compliance with the H-2A law. The ethnographic case studies indicated that in one county the local workers were forced to look for other employment owing to a disparity in working conditions intentionally favoring the seasonal farmworkers. In another county, lower

ENHANCING THE POLICY PROCESS THROUGH THE USE OF ETHNOGRAPHY AND OTHER STUDY  
FRAMEWORKS: A MIXED-METHOD STRATEGY

rates of H-2A farmworkers were not the result of good working conditions favoring jobs for U.S. and legal immigrant workers but instead resulted from illegal fieldworkers allegedly being employed. Importantly, the ethnographic case study and participant observer methodology was viewed as essential to the deep understanding acquired on how the law was being carried out.

Although ethnographic methods have had a long history of use in the federal government, efforts to document and analyze how such methods have contributed to program improvement are still needed. Recently, there has been an acknowledgement about how ethnographic studies can inform agency actions and how it can be used to study culture in organizations. As Patton (2005) noted, the organizational culture is often the context for the program culture; thus, organizational effectiveness and program effectiveness are often interdependent.

Using Ethnographic Methods to Illuminate  
Organizational Culture and Transformation

A transformative shift in agency culture can be captured using ethnographic methods. Under study at GAO was the Department of Veterans Affairs (VA) Patient Safety Program introduced in 1999 to fix system flaws that could harm patients (U.S. GAO, 2004c). Reports about unnecessary deaths and expenditures from accidents in hospitals, the attendant media coverage, and public outcry gave rise to a program intended to minimize the catastrophic effects of physician and staff error. The VA program relied on staff openness to report close calls and adverse events. In the past such reports resulted in hospital employees being held responsible for adverse patient outcomes even when mistakes were of a systemic nature. Success in implementing the new safety initiative required a cultural shift from fear of punishment to mutual trust and comfort in reporting adverse events.

The challenge of determining whether an organizational cultural shift supporting the program had occurred was undertaken at four medical facilities. Ethnographic methods (site visits, including in-depth interviews with VA staff, observations of daily routines and meetings, and document review) were used for the purpose of providing knowledge of organizational culture from the perspective of VA's physicians, nurses, and an array of others responsible for patient care. GAO found that progress in program implementation varied at the different facilities, and this stemmed from significant differences in clinicians' familiarity with and participation in the program. Also, while mainly positive, the levels of cultural support for

the program differed among physicians, and at one facility the culture blocked physician participation. Still, three of the four facilities had a supportive culture so that staff were trusting and comfortable about reporting close calls and adverse events.

In its recommendations, GAO encouraged the VA to set goals for increasing staff familiarity with the program's major concepts (close call reporting, confidential reporting program, root cause analysis) and mutual trust, measure goals by facility, and develop interventions when goals are not met. The use of ethnographic methods to study culture change surfaced a process model for how the Patient Safety Program would lead to desired outcomes through (a) clinicians with cultural support for reporting adverse events and close calls, (b) teams that investigate root causes, (c) systems change, and (d) feedback and reward systems to encourage reporting, resulting in (e) patients who are safer.

In addition to the actual VA study, also of interest is a description of how the study team was able to create a sense of "buy-in" by GAO managers by demonstrating how an ethnographic approach can be as valuable as more typical performance audits used in GAO engagements and could, if appropriate, complement and enhance such audits (Goodman, Trainor, & Divorski, 2005). First, critical to accepting their methodology was whether the ethnographic component could meet stringent audit documentation requirements. The multidisciplinary team cross-validated coding definitions, clarified their interpretations through interrater reliability techniques, and provided audit-trail documentation that was reviewed by a referencer independent of the team for quality assurance purposes. Second, time constraints attend congressional requests and these constraints create challenges considering that ethnography is characterized by prolonged engagement at sites. However, rapid ethnographic assessment is gaining popularity, making ethnographic methodology more amenable to policy time frames. Rapid assessment process (RAP) used by a team offers a team-based ethnographic inquiry that relies on iterative data analysis, including data triangulation, and additional data collection to develop a preliminary "emic" understanding of a situation relatively quickly (Beebe, 2001).

A Coming of Age in the 21<sup>st</sup> Century: The Diverse  
Practice of Ethnography

Potential benefits and limitations of a rapid ethnographic assessment in comparison to traditional ethnography for examining organizational culture,

organizational change, and interrelationships among organizational levels is an area ripe for research. Hammersley (2006), in describing current practice, highlights areas of tension and conflict in ethnography in the social sciences, including educational research, notes the movement from the older anthropological model of ethnographic fieldwork to more recent forms that may attempt only a partial study of people's lives over a relatively short time frame (e.g., months rather than years), with part-time participant observation. These changes reflect increasing pressure to demonstrate productivity, shorter contract time-frames, and uses of new technology. Hammersley, while acknowledging the integral place ethnographic case studies have occupied in conjunction with quantitative methodology, in the study of schools in the U. K., cautions that this shift in practice requires conscientious attention to potential problems of sampling, generalization, cyclical variability, and fundamental patterns of change. Nevertheless, technological changes that allow for the use of portable audio and video recording devices and data analysis software packages that have been developed to assist in the analysis of large quantities of data collected are now part of the evaluator's tool chest and contribute to variation in how ethnography is practiced. Bazeley (this volume) describes the use of qualitative data analysis software packages that both facilitate the analysis of qualitative data and have the capacity for combining and integrating qualitative and quantitative data collected in studies that rely on mixed and/or multiple methods.

These new developments are likely to be looked upon with favor in the policy arena that favors receiving relevant information for decision-making in shorter time frames than traditional ethnographic methods can provide. As illustrated, an ethnographic approach has utility in studying cultural changes in organizations; however, ethnography also can be used concomitantly or sequentially with other study frameworks, within a mixed-method strategy, to secure information in complex study environments.

#### Responding to Federal Program Issues: Enhancing Mixed-Method Studies Through Inclusion of Ethnographic Methods

In 2003, GAO examined the range and scope of the use of ethnography in the federal government. Although it was recognized that ethnography has had a long history of application in the federal government, information about the past and present uses of this methodological approach to improve federal programs had not been systematically gathered or analyzed. While not exhaustive, GAO found that in 10 federal departments or agencies that employed ethnographic

methods, the use of ethnography focused on understanding a group's sociocultural life with respect to an important federal program issue (U.S. GAO 2003). Ethnography was a method of choice when the program's operation or outcomes depended on the actions of a definable cultural community. Some of the benefits cited were the collection of new information, increased understanding of issues or problems important to a program, support for conclusions across sites, reliable identification of recurring themes, and the capacity to integrate ethnographic information with economic or other quantitative data. Several study examples used a mixed-method framework linking an ethnographic study to other quantitative data collection strategies to enhance the quality and accuracy of findings.

For example, the Bureau of the Census has used ethnography to conduct alternative enumerations of urban and rural hard-to-count populations. The Census is the premier source of information about U.S. population and is the basis for apportioning seats in the House of Representatives. Yet, impediments to participation resulting in undercounts, particularly among certain racial and ethnic groups, pose continuing challenges. Ethnography, in use by the Bureau since the 1960s, has offered a means by which the Bureau increases its understanding of why certain groups are undercounted so as to allow an alternative enumeration in the sites studied. Through the Statistical Research Division, the Bureau shed light on critical problems of data quality with ethnographic studies that document undercounts and other enumeration errors, factors contributing to them, and recommendations for improvement.

During the 1988 census "dress rehearsal," ethnographic evaluation studies were conducted in five sites documenting the day residence of all persons enumerated and comparing their observed count to the "official" count and later returned to the field to examine discrepancies. In 1990 ethnographic alternative enumerations in 29 rural and urban sites were undertaken by ethnographers with established relationships with the sites. A complete list of all housing and people in the area, recorded behavioral observations about the neighborhoods, and other factors that could prevent complete counts were studied. After matching ethnographers' counts with the Census count from returned Census forms, the ethnographers returned to the field to reconcile the differences. The ethnographic evaluation found that disparities in the accuracy of the counts for low-income minority populations were attributable to irregular housing units missed in the census, residential mobility, limited English proficiency, distrust of government, and ambiguous housing units (e.g., unrelated individuals or households with two or more

## ENHANCING THE POLICY PROCESS THROUGH THE USE OF ETHNOGRAPHY AND OTHER STUDY FRAMEWORKS: A MIXED-METHOD STRATEGY

nuclear families) that made census residency rules difficult to apply. The combination of centralized control and standardized methods complemented by the ethnographers' local knowledge, expertise, and experience improved the accuracy of the estimation.

In another example, the Administration for Children and Families of the Department of Health and Human Services used ethnographic techniques in two national Early Head Start (EHS) evaluation sites to illuminate ways in which the families served accepted or rejected the program's Montessori Intervention. Early Head Start, a national evaluation with a traditional random assignment design with quantitative measures of process and outcome, included ethnographies at two sites to document more fully the program's socio-cultural context, inform the story of program implementation, and detail more fully what the intervention meant to parents and children living in poverty. One ethnographic study included (a) extensive participant-observation in the classrooms and (b) case studies chosen to illuminate patterns seen in the quantitative data on the impact of the program for 12 families. Preliminary results indicated, in contrast to what may have been believed about Montessori before the program's experience, that low-income parents appreciated and valued the Montessori approach and the growing independence and facility in daily routines that they saw the program produce in their children. This was important because typically the Montessori curriculum had been administered to mid-upper class rather than the study population of low-income children and also because the study confirmed the program's theory that families were changed by first changing the children (Spicer, McAllister, & Emde, 2001).

The other ethnographic study was a nested design in which community and policy developments influencing the operation of EHS were explored. In this study, shifts in the program's theory of change were examined by participant observation of the program activities and focus groups with program staff; ethnographic interviews with program families provided information about their program experiences and their understanding of key program contexts. Relationships among community context, program implementation, and family perspectives were clarified. The ethnography provided a greater understanding of family culture and elaborated the program's theory of change in changing community contexts, which resulted in expanded home-visiting services to informal child care providers and partnering with formal care providers to ensure quality and improve access. This was particularly important given the national evaluation had somewhat limited ability to assess the

community context; yet, the EHS programs were required to tailor services to meet community needs. In both cases, the ethnographies provided insights into aspects of program process that were unanticipated in the design of the randomized trial. (A detailed description of methodology and results related to the study can be found in Mathematica, 2002.)

### Concluding Considerations

As noted at the beginning of this article, federal evaluation is taking place in a policy environment that emphasizes evaluation for accountability purposes. Agencies' performance information is examined by OMB through the PART process, reinforcing the government's focus on program results. PART assessments summarize key performance data and findings, and agencies are scored on program purpose and design, strategic planning, program management, and program results (which account for 50% of the overall score). OMB then makes recommendations to improve performance and may request program assessments (e.g., to develop new measures or conduct program evaluations). As previously stated, the OMB guidance specifies a preference for randomized controlled trials when asking for a demonstration of the program's effectiveness but acknowledges that randomized controlled trials may not always be able to be carried out. As Greene (this volume) points out, an evaluator's sociopolitical commitments can influence the type of knowledge produced and the interests served by a particular approach to inquiry. Yet, the political arena has an influence of its own. Amid claims of scarce resources, the policy environment exerts an influence on the evaluation community through the types of evaluations that receive funding, the skills that are valued, and the request for proposals that are generated.

It is likely that, despite a preference for experimental evidence, policy makers use multiple criteria when trying to maximize the potential of policies and programs to ameliorate persistent problems faced by society, assure the accountability of programs, and effectively serve the needs of the citizenry. This article has elaborated on different, yet valued information that ethnography can provide. While it is the case that an ethnographic study can be designed as a mixed-method study, it is also the case that ethnography can serve as a component or can be integrated into a particular study framework that generates primarily quantitative or multiple forms of evidence (such as, a field experiment or longitudinal survey). Several examples have been provided that demonstrate how the contrasting types of evidence



complement each other and provide in some cases more accurate findings, in others more contextualized information, and in yet others explanation or generation of new insights.

The analytic frameworks and methods depicted here are rarely singly applied in an evaluation. Instead, they are arrayed in different configurations depending upon theories about the program and problems the program is intended to address, the particular questions that guide the evaluation, the skill set and resources of those conducting the study, the user needs for timely information, and other factors. As a result, Datta (1997) wisely advocated a pragmatic approach to the selection of design frameworks and incorporation of mixed-methods that considers the practical consequences of such decisions given the demands, opportunities, and constraints that the situation may place on carrying out a quality evaluation, optimizing resources, and assuring timely, useful findings. This is the approach recommended here as well.

The lead editors for this article were R. Burke Johnson and Anthony J. Onwuegbuzie.

#### References

- Beebe, J. (2001). *Rapid assessment process: An introduction*. Walnut Creek, CA: Alta Mira Press.
- Bernholz, E., et al. (2005). *Evaluation dialogue between OMB staff and federal evaluation leaders: Digging a bit deeper into evaluation science*. Washington, DC. Retrieved June 23, 2006, from <http://www.fedeval.net>
- Boruch, R. F. (1997). *Randomized experiments for planning and evaluation*. Thousand Oaks, CA: Sage.
- Boruch, R. F. (2005). Preface: Better evaluation for evidence-based policy: Place randomized trials in education, criminology, welfare, and health. *Annals of the American Academy of Political and Social Science*, 599, 6–18.
- Brass, C. T., Nunez-Neto, B., & Williams, E. D. (2006). *Congress and program evaluation: An overview of randomized controlled trials (RCTs) and related issues* (Congressional Research Service Report Order No. RL33301). Washington, DC. Available at [http://www.opencrs.com/rpts/RL33301\\_2006\\_0307.pdf](http://www.opencrs.com/rpts/RL33301_2006_0307.pdf)
- Caracelli, V. J., & Greene, J. C. (1993). Data analysis strategies for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 15, 195-207.
- Caracelli, V. J., & Greene, J. C. (1997). Crafting mixed-method evaluation designs. In J.C. Greene & V. J. Caracelli (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms* (pp. 19-32). San Francisco: Jossey-Bass.
- Chatterji, M. (2004). Evidence on “what works”: An argument for extended-term mixed-method (ETMM) evaluation designs. *Educational Researcher*, 33(9), 3-13.
- Chelimsky, E. (1997). The coming transformations in evaluation. In E. Chelimsky & W. Shadish (Eds.), *Evaluation for the 21st century* (pp. 1-26). Thousand Oaks, CA: Sage.
- Cook, T. (2001). A critical appraisal of the case against using experiments to assess school (or community) effects. *Education Next*. Retrieved June 23, 2006, from <http://www.educationnext.org/unabridged/20013>
- Cronbach, L. J. (1982). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.
- Datta, L-E. (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.
- Donaldson, S. I., & Christie, C. A. (2004). *The 2004 Claremont debate: Lipsey vs. Scriven: Determining causality in program evaluation and applied research: Should experimental evidence by the gold standard?* Retrieved June 23, 2006, from [http://www.cgu.edu/include/SBOS\\_2004\\_Debate.pdf](http://www.cgu.edu/include/SBOS_2004_Debate.pdf) (See also, JMDE, 1556-8180, NO. 3, Oct. 2005)
- Droitcour, J. (1997). Cross-design synthesis: Concept and application. In E. Chelimsky & W. R. Shadish (Eds.), *Evaluation for the 21st century: A handbook* (pp. 360–372). Thousand Oaks, CA: Sage.
- Droitcour, J., & Kovar, M. G. (in press). Multiple threats to the validity of randomized studies. In P. Brandon & N. Smith (Eds.), *Fundamental issues in evaluation*. New York: Guilford.
- Fetterman, D. M. (1982). Ibsen’s baths: Reactivity and insensitivity (a misapplication of the treatment-control design in a national evaluation). *Educational Evaluation and Policy Analysis*, 4, 261–279.
- Fry, G., Chantavanich, S., & Chantavanich, A. (1981). Merging quantitative and qualitative research techniques: Toward a new research paradigm.

ENHANCING THE POLICY PROCESS THROUGH THE USE OF ETHNOGRAPHY AND OTHER STUDY  
FRAMEWORKS: A MIXED-METHOD STRATEGY

- Anthropology and Education Quarterly*, 12, 145-158.
- Goodman, C., Trainor, B., & Divorski, S. (2005). Using ethnographic methods to evaluate the Department of Veterans Affairs Patient Safety Program. *NAPA Bulletin*, 24(1), 57-70.
- Greene, J. C. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review*, 12, 91-116.
- 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.
- Hammersley, M. (2006). Ethnography: Problems and prospects. *Ethnography and Education*, 1(1), 3-14. Oxford, UK: Routledge, Taylor & Francis.
- Jacob, E. (1982). Combining ethnographic and quantitative approaches: Suggestions and examples from a study on Puerto Rico. In P. Gilmore & A. A. Glatthorn (Eds.), *Children in and out of school: Ethnography and education* (pp. 124-147). Washington, DC: Center for Applied Linguistics.
- Kirkhart, K. (2000). Reconceptualizing evaluation use: An integrated theory of influence. In V.J. Caracelli & H. Preskill (Eds.), *The expanding scope of evaluation use: New directions for evaluation*, No. 88 (pp. 5-23). San Francisco, CA: Jossey-Bass.
- Lane, J. (2006, May 17). *Optimizing the use of microdata: A new perspective on confidentiality and access* (abstract). Retrieved June 23, 2006 from <http://www.scs.gmu.edu/~wss/Sem2006.html>
- Mathematica. (2002). *Making a difference in the lives of infants and toddlers and their families: The impacts of early head start. Vol. I. Final technical report*. Washington, DC: Department of Health and Human Services, Administration on Children, Youth, and Families. Retrieved June 23, 2006, from [http://www.acf.hhs.gov/programs/opre/ehs/ehs\\_resrch/reports/impacts\\_voll/impacts\\_voll\\_title.html](http://www.acf.hhs.gov/programs/opre/ehs/ehs_resrch/reports/impacts_voll/impacts_voll_title.html) or <http://www.mathematica-mpr.com/earlycare/ehstoc.asp>
- Maxwell, J. A., Bashook, P. G., & Sandlow, C. J. (1986). Combining ethnographic and experimental methods in educational evaluation. In D. M. Fetterman & M. A. Pitman (Eds.), *Educational evaluation: Ethnography in theory, practice, and politics* (pp. 121-143). Beverly Hills, CA: Sage.
- Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 189-208). Thousand Oaks, CA: Sage.
- Office of Management and Budget. (2004). *What constitutes strong evidence of a program's effectiveness*. Retrieved June 23, 2006, from [http://www.whitehouse.gov/omb/part/2004\\_program\\_eval.pdf](http://www.whitehouse.gov/omb/part/2004_program_eval.pdf)
- Patton, M. Q. (2005). The view from evaluation. *NAPA Bulletin*, 24(1), 31-40.
- Spicer, P., McAllister, C. L., & Emde, R. (2001). Ethnography and the Early Head Start evaluation: Contributions from local research to understanding program processes. In Administration for Children and Families (Ed.), *Building their futures: How Early Head Start programs are enhancing the lives of infants and toddlers in low-income families, Vol. II: Technical report, appendixes* (pp. A.77-A.80). Washington, DC: U.S. Department of Health and Human Services.
- U.S. Government Accountability Office. (1994, December). *Managing for results: State experiences provide insights for federal management reforms* (Publication No. GAO/GGD-95-22). Washington, DC: Author.
- U.S. Government Accountability Office. (1998a, January). *Managing for results: The statutory framework for performance-based management and accountability* (Publication No. GAO/GGD/AIMD-98-52). Washington, DC: Author.
- U.S. Government Accountability Office. (1998b, October). *The H-2A program: Protections for the U. S. farmworkers* (Publication No. PEMD-89-3). Washington, DC: Author.
- U.S. Government Accountability Office. (2001, April). *Record linkage and privacy: Issues increasing new federal research and statistical information* (Publication No. GAO-01-126SP). Washington, DC: Author.
- U.S. Government Accountability Office. (2003, March). *Federal programs: Ethnographic studies can inform agencies' actions* (Publication No. GAO-03-455). Washington, DC: Author.
- U.S. Government Accountability Office. (2004a, January). *Performance budgeting: Observations on the use of OMB's program assessment rating tool for the fiscal year 2004*

- budget* (Publication No. GAO-04-174). Washington, DC: Author.
- U.S. Government Accountability Office. (2004b, March). *Results-oriented government: GPRA has established a solid foundation for achieving greater results* (Publication No. GAO-04-38). Washington, DC: Author.
- U.S. Government Accountability Office. (2004c, December). *VA Patient Safety Program: A cultural perspective at four medical facilities* (Publication No. GAO-05-83). Washington, DC: Author.
- U.S. Government Accountability Office. (2005a, February). *21<sup>st</sup> century challenges: Reexamining the base of government* (Publication No. GAO-05-325SP). Washington, DC: Author.
- U.S. Government Accountability Office. (2005b, September). *Hurricane Katrina: Providing oversight of the nation's preparedness, response, and recovery activities* (Publication No. GAO-05-1053T). Washington, DC: Author.
- U.S. Government Accountability Office. (2005c, October). *Performance budgeting: PART focuses attention on program performance, but more can be done to engage congress* (Publication No. GAO-06-28). Washington, DC: Author.
- U.S. Government Accountability Office. (2005d, October). *Program evaluation: OMB's PART reviews increased agencies' attention to improving evidence of program results* (Publication No. GAO-06-67). Washington, DC: Author.

#### Notes

<sup>1</sup> The opinions in this article are the author's and should not be construed as the policy or position of the U.S. Government Accountability Office.

## Toward a Methodology of Mixed Methods Social Inquiry

Jennifer C. Greene

*University of Illinois, Urbana-Champaign*

*The development of a methodology for inquiry into human affairs requires consideration of four interlocking, but nonetheless conceptually distinct domains. Each domain contains a set of issues that are relevant and important to the practice of social inquiry. Collectively, the domains present a justificatory framework and a set of practical guidelines for a given approach to social inquiry. The burgeoning literature in mixed methods approaches to social inquiry engages all four domains, but somewhat unevenly. This is also so for the papers presented in this special issue of Research in the Schools. This concluding article first offers a brief portrait of the requisite domains for a social inquiry methodology, and then connects each of the papers in this special issue to these domains. Comments on priority future directions for the continued development of a methodology of mixed methods social inquiry conclude the discussion.*

### What Constitutes a Methodology for Social Inquiry?

A methodology for social inquiry engages four domains of issues and assumptions: philosophical assumptions and stances, inquiry logics, guidelines for practice, and sociopolitical commitments in science. Each domain is briefly described below.

#### *Domain 1 – Philosophical Assumptions and Stances*

A social science methodology is importantly rooted in issues that are the substance of the philosophy of science, in particular, assumptions about the nature of the social world (ontology) and about the nature of warranted social knowledge (epistemology). This domain also includes stances regarding related issues, such as objectivity and subjectivity, the role of context and contingency in social knowing, and the relationship between the knower and the known. In addition to these traditionally paradigmatic issues, this domain encompasses broader facets of an inquirer's own mental model (Phillips, 1996; Smith, 1997), such as value commitments and the perspectives and core constructs of particular disciplines, for example, "disequilibrium" as a catalyst for growth in human development and "maximization of satisfaction" as the fulcrum of consumer decision making in economics.

---

Correspondence should be addressed to Jennifer C. Greene, College of Education at the University of Illinois at Urbana-Champaign, 230C Education Building, 1310 S. 6<sup>th</sup> St., MC 708, Champaign, IL 61820. Email: jcgreene@uiuc.edu.

---

Domain 1 thus guides the inquirer's gaze to look at particular things in particular ways and offers appropriate philosophical and theoretical justification for this way of seeing, observing, and interpreting.

#### *Domain 2 – Inquiry Logics*

Domain 2 constitutes what is commonly called "methodology" in social science. For a given approach to social inquiry, this domain identifies appropriate inquiry purposes and questions, broad inquiry strategies and designs, sampling preferences and logic, criteria of quality for both methodology and inference, and defensible forms of writing and reporting. The role and location of the inquirer in the study is also delineated in this domain. Further, this domain presents logics of justification for each of these components of social inquiry and especially for their interconnections. A strong inquiry logic is substantiated by coherence and connection among the constituent parts. The separate parts need to fit together and work together to enable – from the perspective of a given inquiry approach – defensible data gathering, analysis, and interpretation.

Domain 2 thus structures the inquirer's gaze, so that what is important to see (as defined in Domain 1) is observed, recorded, and understood or explained in defensible ways. Domain 2 offers a kind of geographic information system (GIS) positioning of the inquirer in the inquiry context and also offers navigational tools that substantially direct the inquirer's journey in that context.

*Domain 3 – Guidelines for Practice*

Domain 3 offers specific guidelines for inquiry practice. The philosophical assumptions and logics of inquiry that comprise domains one and two are translated into particular inquiry steps and procedures in Domain 3. Domain 3 is the “how to” of social science inquiry. This includes, for example, alternative inquiry designs, sampling strategies, and analysis techniques that meet the broad parameters specified in Domain 2. In Domain 3 are also located the specific methods of data gathering, analysis, interpretation, and reporting, for example, alternative interviewing techniques (e.g., life history interviewing, group interviewing) and various statistical procedures (e.g., simultaneous regression, hierarchical linear modeling).

Domain 3 thus provides the specific tools and procedures – the water bottle, hiking boots, and trail map – needed to enact the journey framed by Domain 1 and mapped by Domain 2.

*Domain 4 – Sociopolitical Commitments*

Finally, in Domain 4, the location of the inquiry in society is articulated and defended. Whose interests should be served by this particular approach to social inquiry, and why? Where is this inquiry located in society? Does the study contribute to collective theoretical knowledge; is it a “knowledge producer”? Does it advise governmental decision makers? Is the study located in a protected space, separate and apart from the political fray? Or is it located in the midst of contestation, in a position of social critique or advocacy for particular interests and positions? While Domain 4 is clearly not independent from the other domains, especially Domain 1, the role of social science in society is a distinctive issue.

In the past, science of all kinds was assumed to serve a knowledge generation, testing, and refinement role. The mostly uncontested purpose of science was to generate knowledge that could then be applied to the improvement of society, in material, economic, or other ways. Yet, today, there are recognized inquiry paradigms for social science – notably, various forms of feminism, critical race theory, and LatCrit – that themselves embody distinctive sociopolitical stances. Feminisms privilege the well being of women, and critical race theory occupies a stance of social critique about institutional racism in American society. For other paradigms, sociopolitical location is circumscribed but not completely defined. For example, an interpretivist paradigm acknowledges the multiplicity and contextuality of social knowing and thus inherently values multiple perspectives, while a post-positivist paradigm seeks generalizable causal knowledge and thus inherently privileges cross-context recurring regularities in human action.

Domain 4 thus importantly directs the inquirer’s journey toward a particular destination, as it identifies

priority roles for social science in society and provides values-based rationales and meanings for the practice of social inquiry. While values are present in all four domains, they are proclaimed in Domain 4.

*Interconnections, Coherence, and Persuasiveness*

A methodology for social inquiry gains credibility and persuasiveness when all of these domains act in concert with one another, when their interlocking connections are smooth and well oiled, when the overall presentation is strong, coherent, well articulated and thus persuasive.

An experimental approach to social inquiry is one example of a strong, coherent foundation for inquiry with tightly interconnected logics of justification, positioning, procedures, and rationales. Experimentalism is rooted in a post-positivist philosophy (Domain 1), which espouses a realist view of the social world and an objective, though fallible, stance on social knowledge. The primary purpose of science in society in the experimental tradition is to develop, test, and refine theory, particularly causal and explanatory theory in the Humean tradition (Domain 4). So, priority questions for social inquiry are casual questions about the effects or outcomes of a particular human action, experience, or social intervention. And the randomized experiment is viewed as the best methodology for testing causal hypotheses (Domain 2), as experimentation helps control for many rival hypotheses in the real world and certain biases and errors in human judgment. Experimentation today is well supported by sophisticated procedures of randomization, statistical control, and statistical analyses (Domain 3).

The interpretive case study (Stake, 1995) is another example of a well articulated social inquiry tradition. Rooted in an interpretivist philosophy, the case study approach assumes that the social world, and human actions therein, are guided in part by social constructions of meaning, and so social knowledge is contextual, dynamic, and pluralistic (Domain 1). Interpretive case studies seek not generalizable causal explanations but contextual understanding of the meaningfulness in human experience (Domain 4). A case study methodology honors these assumptions of contextuality and meaning and guides the inquirer to construct and re-interpret an inside or emic portrait of meaningfulness in that context (Domain 2). Case study inquiry itself is well supported today by techniques of purposeful sampling and methods of qualitative data gathering, and by attention to inquirer reflexivity about the presence of the “self” in the study (Domain 3).

Similar portraits can be offered for other well developed methodological traditions, including action research, participatory action research, survey research, and narrative inquiry. A methodology – or perhaps

multiple methodologies – for mixed methods social inquiry is still in the developmental stages.

#### Development of a Methodology for Mixed Methods Social Inquiry

So, how do the articles in this special issue take up the various challenges of developing an overall methodology for mixed methods social inquiry? What domains are represented in this set of articles? And how are these authors engaging with the distinct challenges of each domain?

#### *The Busiest Site of Development is Domain 2*

The preponderance of developmental work represented in this collection of papers on a mixed methods approach to social inquiry is in Domain 2, the site for identification and justification of the broad methodological framework for a given tradition. Included in this domain are issues of inquiry purpose and audience, priority inquiry questions, broad inquiry designs, sampling preferences and logic, criteria for judging quality of both method and inference, defensible forms of writing and reporting, and inquirer role. In the remainder of this section, I discuss the articles that I interpret to be primarily discussions of Domain 2 issues. These articles address issues of central importance to Domain 2: design, research synthesis, and validity.

Robert K. Yin's article makes three primary arguments, all of which pertain to issues of inquiry design. Specifically, Yin's argument engages three continuing issues in mixed methods design. First, should methods be mixed within a given study or only across studies (e.g., Morse, 2003)? Here Yin argues that the essence of a mixed methods approach is its contributions to *convergence* in inquiry findings (as in a triangulation design), which only makes sense when methods are mixed *within a single study*. Mixing methods across studies, Yin argues, "resemble[s] well-trodden paths involving research syntheses (e.g., Cooper & Hedges, 1994), meta-analyses (e.g., Lipsey & Wilson, 2001)," (p. 41) or similar "aggregative procedures" and should therefore not be considered mixed methods inquiry. Second, what kinds of methodological mixes should be included in a definition of mixed methods inquiry? Yin argues that mixes should not be restricted to different qualitative and quantitative methods, but should also include different methods within a given tradition (as in interviewing and observation). This issue is related to discussions of the differences between mixing methods and mixing models, offered most thoughtfully by Tashakkori and Teddlie (1998) and Teddlie and Tashakkori (2003; this volume). Third, when and how should methods be integrated in a study? Here, Yin argues for integration throughout the course of a study,

claiming "the more that a single study integrates mixed methods ... the more that mixed methods research, as opposed to multiple studies, is taking place" (p. 42). Some of the specifics of Yin's discussion on this third point are also related to the "how to" of social inquiry or Domain 3.

The article by Charles Teddlie and Abbas Tashakkori offers another iteration in their thoughtful, stimulating, and useful developmental work on conceptualizing mixed methods design in terms of a typology that is defined and organized around critical dimensions of conceptual difference and practical relevance. Their article engages issues of design typologies in some depth, including a supporting argument for the usefulness of design typologies, an argument for the dimensions included *and* excluded in this typology (drawing on relatively extensive work on design dimensions of importance in mixed methods studies), and an elaboration and illustration of each of the designs presented. Interestingly, this typology is presented for social inquiry more broadly and thus locates mixed methods work within the general domain of inquiry designs. This work is central to Domain 2 as a site of development of a mixed methods approach to social inquiry.

The article by Margarete Sandelowski, Corrine I. Voils, and Julie Barroso also engages a centrally important issue in Domain 2, that of research syntheses involving studies representing different methodological traditions. The work draws on the mixed methods literature as well as on the large body of literature on research review and synthesis. "Mixed research synthesis is our name for the type of systematic review aimed at the integration of results from both qualitative and quantitative studies in a shared domain of empirical research," (p. 29) note these authors. Their work then engages in some depth and endeavors to resolve critical challenges generated by the methodological diversity within and between qualitative and quantitative inquiry traditions. These challenges include contextual challenges, such as the current privileging of "evidence-based" or otherwise quantitative inquiry results in policy circles and the frequent cooptation of qualitative inquiry in mixed methods contexts. These challenges are also intrinsically methodological, including different definitions of inquiry quality in different traditions, different conceptualizations of human phenomena in different traditions, and continuing debate about the feasibility of or sensibility of mixing epistemological paradigms while mixing methods. The response to these challenges offered by these authors is to mirror in some important ways the thinking of theorists regarding mixing within a single study. That is, just as for single mixed methods studies there are component and integrated designs (Caracelli & Greene, 1997), sequential and concurrent designs (Creswell, Plano

Clark, Gutmann, & Hanson, 2003; Greene, Caracelli, & Graham, 1989), and conversion designs (Teddlie & Tashakkori, this volume). Sandelowski et al. offer three designs for mixed methods research synthesis: the segregated, integrated, and contingent designs. This innovative work represents an important step forward in the contemporary mixed methods conversation taking place across *difference*, and it engages difference in this conversation respectfully and generatively.

The fourth Domain 2 article in this volume engages another issue of central importance to a methodological framework and justification for an inquiry approach – the formidable but critical issue of validity. Like Sandelowski et al., Anthony Onwuegbuzie and Burke Johnson conceptually engage an underdeveloped but critically important issue of methodology, namely, what constitutes quality in mixed methods inquiry, specifically quality related to inferential claims of truth, credibility, or more generically validity. Their argument is rooted in their own particular conceptualization of mixed methods inquiry as inquiry that “involves combining complementary strengths and nonoverlapping weaknesses of quantitative and qualitative research,” (p. 60) and validity in this article thus refers to both methodological defensibility as well as defensibility of inferences or conclusions. Building on Teddlie and Tashakkori’s notion of *inference quality* (2003), Onwuegbuzie and Johnson offer the concept of *legitimation* as a conceptual framework for mixed methods validity. They position the legitimation concept as directly engaging the mixed methods challenge of integrating data and interpretations from very different frameworks, stances, methods, samples, and analyses. These authors then continue by positing nine distinct types of legitimation, each referring to a different strand of methodology (e.g., sampling) or to a different type of mixed methods design (e.g., sequential or concurrent). The article further offers brief windows into other domains; the brief discussion on pragmatism is related to the philosophical issues in Domain 1 and the notion of “political legitimation” represents a tickle in Domain 4. Overall, this article is rich with generative ideas and possibilities and, like the Sandelowski et al. article, represents an important step forward in the development of a methodology for mixed methods social inquiry.

#### *Domain 3 is the Site of Some Creative Ideas and Insights*

Two of the remaining articles in this special issue offer creative ideas related primarily to Domain 3, or how to conduct mixed methods social inquiry. These are the articles by Pat Bazeley and Huey Chen.

Pat Bazeley’s article addresses the conceptual and practical challenges of *integrated* mixed methods data

analysis – that is, analysis of numbers and words – and, as such, spans issues from both domains two and three. The specific focus of Bazeley’s work is on computer software that enables integrated analyses. Yet, her typology of types of integrated data analysis – combined, converted, and blended – and her exceptionally useful examples extend this practical advice to meaningful conceptual ideas and concepts. Bazeley strives in her article to “imagine and envision what might be possible – to tread new paths,” (p. 65) and she accomplishes in this ambition very well.

Huey T. Chen presents an argument for using a mixed methods methodology for a theory-driven approach to program evaluation. Theory-driven evaluation is framed by the substantive issues at hand, for example, an intervention strategy to abate substance abuse among youth or an economic incentive to work among welfare recipients. This is in contrast to the customary privileging of methodology in evaluation studies. Given that substantive issues are inherently complex, a mixed methods approach is sensible for an evaluation oriented around issues of substance. Within this discussion, Chen offers four interesting strategies for mixing methods in a program theory-driven evaluation context: switch, complementarity, contextual overlay, and triangulation assessment. These methodological ideas, as framed within a specific site of application, might contribute usefully to the continued development and empirical critique of mixed methods strategies.

#### *A Modest Foray Into Domain 1*

The article by John Creswell, Ron Shope, Vicki L. Plano Clark, and Denise O. Green is intended to offer a counter-argument to recent critiques from leading qualitative inquirers that mixed methods inquiry represents a cooptation of the basic premises, stances, and ambitions of interpretivist, constructivist qualitative inquiry. Specifically, the critiques maintain that mixed methods inquiry relegates qualitative methods to a secondary or auxiliary status and does not honor or preserve the distinctive philosophical and value stances of qualitative traditions. The Creswell et al. article is thus located within the philosophical arguments that comprise Domain 1. And the authors construct their counter-argument by (a) citing the pronouncements of selected qualitative inquirers that mixed methods inquiry makes ample room for qualitative logics of justification, (b) offering some empirical examples of mixed methods inquiry in which qualitative methods played an important role, and (c) citing a mixed methods design typology in which qualitative methods can be an important, even dominant framework for an empirical study.

The authors further focus on particular dimensions of interpretive frameworks, namely, empowering participants and recognizing the historical and socio-

political context of the inquiry. While important, these are not the fundamental philosophical tenets of interpretive or constructivist ways of knowing; not addressed are issues of ontology or epistemology nor how different stances on reality and social knowledge can co-exist in the same study. So, while offering a brief engagement with the philosophical concerns of Domain 1, this article does not venture very far into the challenging terrain of philosophical frames for making meaning and knowledge in social inquiry.

#### *A Meaningful Engagement with Domain 4*

The final article in this collection, by Valerie Caracelli, offers a window into the intentional mixing of methods for better understanding, defined as the provision of political voice. This article thus engages with the sociopolitical commitments of Domain 4. In the article, Caracelli develops an argument for the importance of including ethnography in federal policy-relevant evaluation. She first observes that the current accountability demands and practices within the federal government narrowly constrain the kinds of evaluation studies favored and especially the kinds of information generated in such studies. Notably, neither information on program processes and participant experiences nor information on important features of program contexts is usually collected as part of a randomized experiment. Yet, this information is required for a "fully responsive" evaluation that generates "deep understanding" of particular program experiences and outcomes in particular contexts.

Moreover, argues Caracelli, ethnography can complement experimental or other quantitative methodologies in important ways by providing valuable information relevant to federal policy making that extends beyond outcome measurement and attribution. Ethnographic information can illuminate organizational culture and transformation, respond more fully to the multiple criteria actually used by policy makers in their decisions, and especially can give voice to the perspectives, experiences, and realities of program participants and of marginalized communities. Examples of the latter provided by Caracelli include a GAO study on immigration, efforts by the Census to address the "undercounting" of marginal populations such as the homeless, and studies of Early Head Start in low-income communities. Caracelli's recognition of the politics of method and the potential of mixed methods social inquiry to engage these politics in service of voice and equity are highly commendable.

#### Future Directions

Continued development in all domains of a mixed methods methodology for social inquiry is needed, perhaps those especially under-represented in this

collection of articles. While Domain 2 represents a critically important site for further development, other sites are less well represented in the contemporary mixed methods conversation and yet also present important contexts and issues for creative and thoughtful developmental work.

In the groundbreaking *Handbook of Mixed Methods*, Teddlie and Tashakkori (2003) cited six important issues in the field. These six issues map well onto the four domains of methodological development presented in this discussion. Relevant to Domain 1, Teddlie and Tashakkori (2003) cited as a key issue in the field, "the paradigmatic foundations for mixed methods research." For Domain 2, these authors cited "design issues" and "issues in drawing inferences," well represented in this volume by the Teddlie and Tashakkori and the Onwuegbuzie and Johnson articles, respectively. Connected to Domain 3, the *Handbook* editors note the importance of "the logistics of conducting mixed methods research" and in Domain 4, "the utility of mixed methods research (why do we do it?)" (The final issue cited by Teddlie and Tashakkori refers to language and nomenclature.) I would echo Teddlie and Tashakkori's developmental priorities, but again encourage participants in the mixed methods conversation to consider needed developments in all domains.

My additional thoughts about important future directions for the development and advancement of a methodology for mixed methods social inquiry include the following. *First*, I encourage contributors to this developmental work to be as welcoming of divergence and dissonance as of convergence and consonance. One vitally important role for mixed methods social inquiry is to trouble taken-for-granted understandings or assumed common meanings of constructs by incorporating a diversity of perspectives, voices, values and stances. In this role, mixed methods inquiry honors complexity alongside diversity and difference, and thereby resists simplification of inherently contextual and complex human phenomena. Convergence and consonance in the service of stronger validity are indeed important contributions of mixed methods inquiry, but so are divergence, dissonance, and difference. All are valuable and important.

*Second*, and related to the first point, I encourage participants in the mixed methods conversation to keep the debate about the "paradigmatic foundations for mixed methods research" alive and lively. As many authors have suggested (including many of the authors in this volume), there may well be an alternative paradigm that offers a sensible and substantive positioning and justification for mixed methods inquiry. And critical realism (Maxwell, 2004) or pragmatism (Biesta & Burbules, 2003) may indeed be strong candidates for such an alternative paradigm. But, for this argument to be made persuasively, further



work is needed on just how such a paradigm shapes and guides mixed methods practice. For example, where do the consequential, actionable assumptions about social knowledge that are advanced in most pragmatic philosophies show up in practice? What does such knowledge look like, and how is it validated? Or perhaps there is not just one paradigm that meaningfully justifies and guides mixed methods social inquiry. Perhaps a distinctive characteristic of a mixed methods methodology is its paradigmatic pluralism.

*Third*, it is important for the vitality and viability of a mixed methods methodology for social inquiry that it be relevant and useful to multiple domains of human activity. We therefore need participants and perspectives from all corners of the social science community, including traditional academic disciplines of psychology, sociology, and even economics, as well as applied fields like nursing, social work, education, and organizational development. Perhaps current participants in this developmental effort can invite colleagues from other fields to join them.

*Fourth* and finally, the development of a mixed methods methodology will continue to be dynamic and open to good ideas and fresh insights if we continue to honor both theory and practice. Some of the most generative conceptual ideas about mixing methods have come from field trials of mixed methods ideas or analytic reviews of mixed methods practice, most recently by Katrin Niglas (2004). It is not enough to think well; we must also demonstrate the value and importance of a mixed methods way of thinking in our practice.

The lead editors for this article were R. Burke Johnson and Anthony J. Onwuegbuzie.

#### References

- Bazeley, P. (2006). The contribution of computer software to integrating qualitative and quantitative data analyses. *Research in the Schools, 13*(1), 64-74.
- Biesta, G. J. J., & Burbules, N. C. (2003). *Pragmatism and educational research*. Lanham, MD: Rowman & Littlefield Publishers.
- Caracelli, V. J. (2006). Enhancing the policy process through the use of ethnography and other study frameworks: A mixed-method strategy. *Research in the Schools, 13*(1), 84-92.
- Caracelli, V. J., & Greene, J. C. (1997). Crafting mixed-method evaluation designs. In J.C. Greene & V. J. Caracelli (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms* (pp. 19-32). San Francisco: Jossey-Bass.
- Chen, H. T. (2006). A theory-driven evaluation perspective on mixed methods research. *Research in the Schools, 13*(1), 75-83.
- Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209-240). Thousand Oaks, CA: Sage.
- Creswell, J. W., Shope, R., Plano Clark, V. L., & Green, D. O. (2006). How interpretive qualitative research extends mixed methods research. *Research in the Schools, 13*(1), 1-11.
- 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.
- Maxwell, J. A. (2004). *Realism as a stance for mixed methods research*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 189-208). Thousand Oaks, CA: Sage.
- Niglas, K. (2004). *The combined use of qualitative and quantitative methods in educational research*. Dissertation, Faculty of Educational Sciences Tallinn Pedagogical University, Tallinn, Estonia.
- Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools, 13*(1), 48-63.
- Phillips, D. C. (1996). Philosophical perspectives. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 1005-1019). Old Tappan, NJ: Macmillan.
- Sandelowski, M., Voils, C. I., & Barroso, J. (2006). Defining and designing mixed research synthesis studies. *Research in the Schools, 13*(1), 29-40.
- Smith, M. L. (1997). Mixing and matching: Methods and models. In J. C. Greene & V. J. Caracelli (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms. New Directions for Evaluation 74* (pp. 73-85). San Francisco: Jossey-Bass.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. 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 and C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 3-50). Thousand Oaks, CA: Sage.
- Teddlie, C., & Tashakkori, A. (2006). A general typology of research designs featuring mixed methods. *Research in the Schools, 13*(1), 12-28.
- Yin, R. K. (2006). Mixed methods research: Are methods genuinely integrated or merely parallel? *Research in the Schools, 13*(1), 41-47.