conducting social inquiry that meaningfully engages with difference and that is thus positioned in service to the public good, toward a noble vision of a pluralistic society characterized not by radical disparities in power and privilege, but by tolerance, understanding, and acceptance.

LOOKING AHEAD

Given this view of a mixed methods way of thinking, mixed methods inquiry is defined in this book as the planned and intentional incorporation of multiple mental models—with their diverse constituent methodological stances, epistemological understandings, disciplinary perspectives, and habits of mind and experience—into the same inquiry space for purposes of generatively engaging with difference toward better understanding of the phenomena being studied. Mixed methods inquiry thus defined has valuable potential to contribute to better understanding that is distinctively marked by greater tolerance, acceptance, and respect for difference.

Chapter Three presents some of the history of the contemporary interest in mixed methods social inquiry, including the "great quantitative-qualitative debate" of the 1960s, 1970s, and 1980s. Yet in this historical discussion, and throughout the remainder of this volume, the terms "qualitative" and "quantitative" will be reserved to refer to methodological traditions, methods, and forms of data, *not* to philosophical paradigms. Paradigms will be labeled with their appropriate monikers; for example, post-positivism and constructivism. This usage is intended to reinforce the importance of attending to philosophical assumptions, disciplinary perspectives, and other substantive components of mental models when conversing about mixing methods. The quantitative and qualitative labels make it too easy to focus on designs, methods, and data alone. They make it too easy to position the conversation at a technical level only, rather than at a level that encompasses issues related to the nature of knowledge, different ways of seeing and knowing, and varied purposes for social inquiry.

CHAPTER 3

THE HISTORICAL ROOTS OF THE CONTEMPORARY MIXED METHODS CONVERSATION

A BIT of time travel into the recent past awaits the traveler in this chapter. The primary purpose of this time travel is to engender a thoughtful understanding of the origins of the contemporary interest in mixed methods social inquiry, under the premise that history is always instructive. Travelers thus will revisit the character of and issues constituting the qualitative-quantitative debate that occupied several decades of ferment in multiple social science communities during the latter half of the twentieth century.

For as long as I can remember, quantitative, qualitative, and mixed-model evaluations have co-existed. In the early days of the U.S. Office of Economic Opportunity, for example, *there were quasi-experimental designs (for instance, both the Westinghouse-Ohio study and the Head Start Longitudinal Study . . . used quasi-experimental designs). There were mixed-model designs, such as the classic Head Start Community Impact Study. [And] there were qualitative studies, such as Greenberg's The Devil Wears Slippery Shoes, a report on the Mississippi Community Action Program. . . . (Datta, 1994, p. 54)*

NOT NEW

The practice of mixing methods in empirical studies is not a new phenomenon, especially in the highly applied domains of social science that are dedicated to understanding and improving human practices in the real world, including education, social psychology, sociology, organizational studies, and evaluation. I have long surmised that methodological openness to new ideas—qualitative inquiry in the 1970s and mixed methods inquiry more recently—is more characteristic of highly applied domains than domains centered on laboratory research, precisely because the complex and messy demands of inquiry in the real world compel acceptance of multiple strategies and tools for understanding. Lois-ellin Datta's (1994) storied reflections on her decades of evaluative inquiry in service to public programs attest to this observation. But the current buzz about mixing methods in social inquiry *is* a more recent phenomenon—one that gradually took root in the 1980s, sprouted a few buds during the 1990s, and then started to blossom at the turn of the century. The current conceptual work in mixed methods inquiry is dedicated to giving some theoretical form to these burgeoning ideas about mixed methods inquiry. Indeed, many contemporary mixed methods concepts arose from reflective scrutiny of early pioneering empirical work (Greene, Caracelli, & Graham, 1989). And indeed, mixed methods "theory" proceeds well only in respectful conversation with practice.

The contemporary interest in mixed methods approaches to social inquiry is a natural and logical development, given the recent history of social scientific thought and practice—a history that spans the latter half of the twentieth century, with roots in prior eras. Understanding the main outlines of this history is essential to a full appreciation of the continuing issues and controversies in mixed methods inquiry. Historical grounding can significantly enrich meaningfulness and understanding. This chapter presents these historical outlines of the contemporary mixed methods conversation. The discussion starts with the seeds of discontent—sown mid-century both in the rarified realms of the philosophy of science as well as in the nitty-gritty trenches of applied social science practice, particularly program evaluation. This discontent shortly (in the 1960s and 1970s) took the form of a full-fledged battle of words and ideas, known then and now as the "great qualitative-quantitative debate" ("great," of course, only to those in the social scientific community who cared about it). The substance of this debate will be only briefly recounted, as there are already numerous excellent reconstructions of the arguments at hand. Some of the character and contours of the debate will also be shared, as those were intellectually challenging and very exciting times. The rapprochement that signaled the end of this great debate will then be discussed, with a particular focus on the logical emergence of rapprochement in the form of mixing methods; that is, as the inclusion of qualitative and quantitative methods in the same study. This idea had appeal

in part because of the fairly long-standing acceptance of methodological triangulation in *both* qualitative and quantitative methodological traditions. The concept of triangulation and its origins in both traditions will be described as part of this history.

So a truce, if not a peace, was declared, although not all at once, not among all combatants, and not forever. And troubled waters remained. If one mixes methods, does that mean one also mixes paradigms? Is such an idea possible, sensible, practical? What are the politics of mixing methods? Is this just a ruse to silence those qualitative upstarts once and for all (Smith and Heshusius, 1986)? Witness the recent reengagement of these issues in the scholarly community in reaction to a reprivileging of quantification, standardization, outcomes, and even experimentation in some arenas by many western governments' wholesale adoption of new managerial practices and accountability-oriented policies during the past two decades (Power, 1999). Further, if methods are mixed, does this threaten the integrity and the quality of each method? And what about all the other paradigms and methodologies that have emerged in recent decades? Can one mix methods with a feminist paradigm? A participatory approach to social inquiry? A post-modern reading of social life as radically contingent and uncertain, best conceptualized as text and cinema? These and other questions that accompanied the emergence of the idea of mixing methods helped to catalyze further interest in and development of this approach to social inquiry.

Now to the brief recounting of the recent history in social scientific methodology that gave birth to the current mixed methods conversation.

THE PHILOSOPHICAL SEEDS OF DISCONTENT

At their beginnings, social sciences in many western societies, including the United States, were modeled after the natural sciences. Like their natural science colleagues, social scientists sought to induce theory from pristine observations and then test this theory under carefully controlled experimental conditions, toward empirically sound and generalizable propositions about how the social world worked. The purpose of this experimentalist social science, that is, was to explain and thereby predict and control the social world, just as physicists explain and endeavor to predict and control the worlds of quarks and quasars.

This way of thinking about social science was implicitly framed by a positivist and later a post-positivist paradigm. Post-positivism retained the ideals of positivism—the assumption that the social world exists independent of our knowledge of it (realism), a commitment to objective methods and to methodological sophistication, and the setting of questions of value outside the perimeter of scientific questions of fact, all in service of causal explanation as universal truth—but post-positivists retained these ideals with more humility, less faith in the power of method, and better acceptance of the inevitable fallibility of human beings as observers than their forbearers had. For example, acknowledging this fallibility, Donald Campbell, an exceptionally eloquent spokesperson for post-positivist thought during the twentieth century, advocated the submission of all social scientific knowledge claims to a "disputatious community of scholars" (Campbell, 1984, p. 44) to see which survive the tests of intellectual challenge, empirical replication, and time. (For further discussion of the philosophical evolution from positivism to post-positivism, see Cook, 1985; Phillips, 1990; and Phillips & Burbules, 2000.) This framing of social science by the tenets of post-positivism was implicit because this framework was neither acknowledged nor articulated, so it could not be challenged. One became a social science investigator mainly by acquiring strong skills in the accepted methods of inquiry of the day. And methods in post-positivism continue to be designed to serve, in large part, to protect the data from the particular predispositions and stances of the inquirer. Given this vitally important role, post-positivist methods grow ever more sophisticated with each advance in knowledge and technology.

Meanwhile, on the heights of erudite thinking about the nature of science, philosophers were fully engaged in conversations and debates about other frameworks for social science and had been since the latter part of the nineteenth century. At that earlier time, the philosopher Wilhelm Dilthey contended that there was a fundamental difference in subject matter between the natural sciences and the social sciences.

> *Whereas physical sciences dealt with a series of inanimate objects that could be seen as existing outside of us (a world of external, objectively knowable facts), [social] sciences focused on the products of the human mind with all its subjectivity, emotions, and values. From this [Dilthey] concluded that since social reality was the result of conscious human intention, it was impossible to separate the interrelationships of what was being investigated and the investigator. There was no objective social reality as such divorced from the people, including investigators, who participated in and interpreted that reality. . . . [So] the investigator of the social world could only attain an understanding of that world through a process of interpretation—one that inevitably involved a hermeneutical method. [Further] the meaning of human expression was context-bound and could not be divorced from context. (Smith & Heshusius, 1986, p. 5)*

Dilthey's interpretive stances for the social sciences stimulated multiple projects in the philosophy of science, as these stances raised considerable challenges to post-positivist assumptions. Instead of objective, realist, and generalizable claims of truth, interpretive knowledge was viewed as inherently and inevitably subjective, contextual, contingent, and value-laden.

By the middle of the twentieth century, these philosophical disputes had made their way to the communities of social science theorists and methodologists, quite a number of whom jumped into the fray. Social scientists' willingness to engage these complicated issues was encouraged by practical difficulties with the methods of post-positivist science (as discussed in the next section) and also, for some, by the general revolutionary temperament of the 1960s and 1970s, at least in the United States. Social movements abounded in that era—movements that sought to end the Vietnam War, shatter the shackles of discrimination and prejudice, and experiment with new social mores and norms. Disrupting the canons of high science, initiating conversations about the politics of method, and envisioning new horizons and possibilities for social science—these were easily embraced by those already wearing multicolored armbands of protest.

As will be elaborated in the discussion that follows on the great debate, these weighty issues about the very nature of social science were engaged both at the level of philosophy or worldview and at the level of methodology and technique. This is partly because many methodologists were quite comfortable engaging with the merits and limitations of standardized surveys compared to contextualized participant observation, but less comfortable engaging with the perennial philosophical debates about objectivism versus relativism or realism versus idealism. In fact, one indelible feature of that era was the heady immersion of many methodologists and practitioners in the abstractions of philosophy and the steep learning curve such immersion entailed. Again, one learned how to be a social scientific inquirer at that time by mastering method, not by pondering paradigms—because the exclusively dominant paradigm remained unspoken.

THE SEEDS OF DISCONTENT IN PRACTICE

The midpoint of the twentieth century in the United States was also the era of the War on Poverty, a massive social welfare campaign mounted by the federal government. Programs supported by considerable federal funding were initiated in nearly all sectors of social well-being—including education, employment, health, welfare, community development—with the collective intention of eradicating poverty in the country once and for all. A number of highly esteemed applied social scientists of that era—sociologists, psychologists, educators, and economists, including esteemed scholars Donald Campbell, Lee Cronbach, Carol Weiss, Peter Rossi, among others—were sufficiently beguiled and challenged by this enormously ambitious attempt at societal change to offer their expertise in studying its effects. Also, some of the federal initiatives were actually developed in tandem with evaluative mandates—perhaps most notably, the Elementary and Secondary Education Act of 1965.

So accompanying the War on Poverty were policy-oriented investigations of the quality and effectiveness of its many innovations and interventions—investigations that used the best and most sophisticated methodologies of that time, which were still the experimentalist methods of the post-positivist paradigm. But as has been well documented (with a near-legendary reputation), experimental methodologies did not, and in fact could not do the job. The need for randomization of potential program participants into control and experimental groups raised substantial ethical issues; even when possible, randomization was vulnerable to contamination. The important requirements of experiments for careful control over and standardization of the conditions under which participants experienced the intervention simply did not work in the real world of social interaction. In particular, defensible experimental research required standardized strategies in tightly controlled contexts, but good practice required constantly adjusting strategies to changing circumstances (Marris and Rein, 1982, p. 206). In programs such as Community Action Agencies this mismatch was clear. Peter Marris and Martin Rein explain, *As soon as the staff, through experimentation or trial and error, discovered a better way of serving trainees they adapted their procedures, methods and techniques accordingly. It was impossible to be inventive, flexible and expedient on the one hand and at the same time do careful, scientific, controlled research on the other (1982, p. 198).*

Critics further argued that confining social inquiry to observable phenomena constrained, indeed even biased research on social programs and problems to the finite ability of research methods to capture such phenomena. Such a requirement also privileged the assessment of program inputs and outcomes, with but scant attention to program processes or experiences—a criticism later dubbed the "black box" approach to evaluation. Cook's post-positivist critical multiplism (1985), presented in Chapter One, is in significant part a response to this failure of the methods of experimental science in real world contexts. He summarizes,

We can see in the attack on [post-]positivist methods a rejection of the primacy of observation over introspection, quantification over understanding, micro-level over macro-level analysis, control over naturalism, theory testing over discovery, and crucial experiments conducted on select parts of nature over more tentative probing of all of nature. (p. 29)

In short, the experimental methods of post-positivist social science were not able to provide sound empirical data on the quality and effectiveness of the War on Poverty programs. In the field of evaluation, the time period following these failed experimental studies of the innovations of the War on Poverty (1970s and 1980s) saw an explosion of new ideas and theories about evaluation, some of which embraced the qualitative methodologies of the newly understood interpretive and constructivist paradigms and, some a bit later, more ideologically oriented paradigms of participation, social action, and social justice.

These seeds of discontent, in the form of serious challenges to standard social science practice, were experienced perhaps most acutely in the policy and program evaluation communities, but were certainly repeated elsewhere. The field of educational research, for instance, was fully engaged in these issues and controversies. The social science methodology pendulum was swinging hard and fast toward qualitative approaches to social inquiry, toward studies of people's contextualized experiences, toward inquiry rendered as narratives. But it wasn't swinging without opposition. Rather, what ensued was fifteen to twenty years of intense and sometime rancorous debate about the relative merits of comfortingly familiar quantitative methodologies compared with the discomforting strangeness of qualitative methodologies—along with their philosophical roots and rationales.

THE GREAT QUALITATIVE-QUANTITATIVE DEBATE

So conventional wisdom was being challenged in both theory and practice. When the assumptions and stances of post-positivism were articulated and held up for scrutiny, they were found wanting by some social scientists of the era. They were found wanting

also because of the practical challenges they encountered on the ground, in the real world outside the laboratory. Experimental methods did not work as well with human beings as they did with crops or chemicals or magnetic forces. Human beings behaved with unpredictable intentionality, whereas crops and chemicals and magnetism appeared to behave in accordance with manipulated conditions. And the contexts inhabited by humans were not nearly as controllable as soil in fields or equipment in laboratories. Yet the assumptions and stances of interpretivism, constructivism, phenomenology, and other paradigms favoring qualitative methodologies appeared radical, even ludicrous to other social scientists—clearly an instance of throwing the baby out with the bathwater. What kinds of explanations and theories could be developed from single intensive case studies? How could empirical results be trusted if derived so subjectively? And what was really accomplished anyway by one interview with one person in one place at one time?

What distinguishes the debate that gained ground in the 1970s was the systematic and self-conscious intrusion of broader philosophical issues into discussions of methods of research. . . . [And] there seem, then, to be two fairly distinct versions of the nature of the difference between quantitative and qualitative research which might usefully be referred to as the "epistemological" and the "technical" accounts. However, there is a tendency for many writers to oscillate between these two versions. (Bryman, 1988, p. 2, 107)

As astutely observed by Bryman, the qualitative-quantitative debate occurred on two main levels—the philosophical and the methodological—even as participants intermingled these two levels while contesting the issues and even as the debate also involved politics and values. A sampling of the key issues—philosophical, methodological, and political—that constituted the debate includes the following. (See also Greene & Henry, 2005.)

Issues at the Philosophical Level

As previewed earlier by Dilthey's ideas, central to the debate were different assumptions about the nature of the social world and the nature of social knowledge. In post-positivism, the social world is assumed to be real; that is, it exists independent of our knowledge of it. And it is assumed to operate much like the physical or natural world; that is, just as plants predictably react to varying environmental conditions of moisture, light, and temperature, so do humans predictably react to varying forces and factors in the external environment. Figuring out just what these forces and factors are, which is most influential, and how they interact is precisely the job of social scientists. In contrast, in most interpretive paradigms, the social world is assumed to be importantly constructed, at least in part. Human beings, unlike plants, act with intentionality. This intentionality is rooted in the meanings that people construct of various phenomena they encounter in their daily lives as they interact with others in varied settings. And it is these constructed meanings that significantly guide and shape human behavior, more so than external forces and factors. Moreover, because different contexts present different constellations of people, interactions, and events, what is meaningful to a given individual or group is, in important measure, context-specific rather than universal.

Post-positivist knowledge claims ideally constitute generalizable causal explanations of observed human phenomena. Post-positivists aim to generate, test, and substantiate theories about human behavior, which comprise empirically warranted propositions about what causes people to do particular things, such as learn well, become a star athlete, or engage in active citizenship or in criminal behavior. Interpretivists aim for contextualized understanding of the meaningfulness of humans' lived experiences. This form of social knowledge is not generalizable nor propositional in form, but rather multiplistic, dynamic, and contingent. Peer tutoring may help one child learn well but not help another, because these two children experience and make meaning of peer relationships in very different ways. Interpretivist knowledge claims are also characteristically textual, narrative, and holistic.

And post-positivists' knowledge claims are warranted by their correspondence with social reality. Truth is attained when theoretical predictions are supported by empirical data. Central to strong knowledge claims, or claims to truth, is excellence in method, so warrants are also constituted by evidence of methodological excellence. Interpretivists' knowledge claims are warranted by the persuasive power of the account. Interpretivists do not claim truth-status for what they come to know, because there are multiple truths or multiple meanings of human experience in any given context. Good method supports persuasive interpretive accounts. But such accounts are at core interpretations—by the person of his or her experience, by the inquirer of the person's interpretation of his or her experience, and then by the reader of the inquirer's interpretation of the person's interpretation of his or her experience (Van Maanen, 1995). These three moments of interpretation encapsulate the nonfoundationalist epistemology that significantly distinguishes interpretive paradigms from the foundational premises of post-positivism. (See J. K. Smith, 1989, for further discussion of this critical difference.)

Paradigma interpretativo

Issues at the Methodological Level

As Bryman noted about this grand debate, intermingled with claims and counterclaims about the nature of the social world, the knowledge we can have about that social world, and thus what is important to know, were charges and countercharges about which set of methods for social inquiry was superior and why. Some of this intermingling was not by design, but some was. In fact, one contested issue in the debate was the necessary coupling of philosophical assumptions with particular methodologies. Do post-positivist beliefs require quantitative methods, and do constructivist commitments mandate the use of qualitative methods? This particular issue is discussed further in chapters Four and Five, as it bears on the role of philosophical assumptions in mixed methods inquiry. Although not yet settled—as few philosophical issues ever are—there emerged some common understandings about the relationship between philosophy and methodology, understandings central to the mixed methods movement and continuing conversation.

The contours of the overall debate at the methodological level are certainly now well known and familiar to the applied social science community. Quantitative methodologies were touted for their perceived superiority as carefully controlled and standardized assessments of human phenomena, as studies of samples carefully and randomly drawn from identified populations, and thus as capable of yielding generalizable claims about the human phenomena under study. Quantitative methods rely on a priori definitions (conceptual and operational) of what is being studied; on a priori designs, methods, and representative samples; and on limits to the number of statistical tests that can be conducted without violating error parameters established for purposes of confidence. Setting up the study in advance fulfills the deductive conditions and expectations of hypothesis testing. Social inquiry, after all, serves the primary purpose of theory testing and refinement. So quantitative methods, properly done, generate propositional findings with confidence, if not total certainty.

In contrast, qualitative methodologies were advanced for their perceived superiority as thoughtful studies of lived human experience, as intense and in-depth studies of a few cases or a few people—purposefully, not randomly selected, and thus as capable of yielding holistic understanding of human behavior in all of its contextuality and complexity. Qualitative methods are emergent and flexible, as they endeavor to be responsive to what is learned as the study proceeds. Qualitative methods center around the perceptive acuity and relational capabilities of the inquirer, as the inquirer is the primary instrument of data generation, analysis, and interpretation. Not setting up the study completely in advance fulfills the inductive, emergent, and contextual challenges of understanding lived experience, which is, after all, the primary purpose of social inquiry. Qualitative methods endeavor to generate this understanding from the perspective of those living the experience, but as necessarily interpreted through the lens of the observer-inquirer.

While these and other contrasting lists of key characteristics of quantitative and qualitative methodologies occupied many pages and hours of the debate, some of the most hotly contested and critical issues concerned this last point: the ways in which and the extent to which the predispositions and stances of the inquirer matter to the quality of inquiry findings. This point encompasses issues related to objectivity and subjectivity, bias and interpretation, independence and engagement, distance and closeness. In quantitative methodological traditions, controlling for bias of all kinds is of vital importance to the quality of the study and is in fact a driving force underlying methodological advancements. In quantitative traditions it is method that protects the data and thus the inquiry findings from the idiosyncrasies of the inquirer. Quantitative inquiry findings are credible only if and when they can be defended as unbiased, objective, independent of the particular stances of the inquirer and of particular theoretical or policy ideas. Quantitative findings tainted by the favored theories or stances of the inquirer are rejected as biased and as unsubstantiated by independent and objective methods and critique.

In qualitative traditions, recalling Dilthey, "it [is] impossible to separate the interrelationships of what was being investigated and the investigator. There [is] no objective social reality as such divorced from the people, including investigators, who participated in and interpreted that reality" (Smith & Heshusius, 1986, p. 5). Distinctive to social inquiry, say the qualitative methodologists, is the inevitable participation of the inquirer in the "social reality" being studied and thus the inevitable presence of the inquirer's own perspectives and understandings in the findings and understandings generated. Human inquiry is inherently subjective and interpretive, they argue. No particular kind or number of sophisticated methods can insulate social knowledge from the particular predispositions of the knower. It is not possible for human inquirers to stand outside their own sociocultural history and location in the world and observe human phenomena with complete impartiality. It is only possible to observe from within one's own historical location; thus human inquiry is inevitably interpretive and inherently subjective. Subjectivity is not bias; rather, it intrinsically defines the very character of human understanding.

As does the intertwinement of "values" with "facts," maintain qualitative methodologists. The fact-value dichotomy long embraced by quantitative methodologists—and its rejection by qualitative methodologists—remains an issue of continuing contestation and debate (House & Howe, 1999; Phillips, 2005). Issues of values are important, argue quantitative methodologists, but not really subject to empirical investigation and so not appropriate for the work of social scientists. Rather, values are best allocated to the domains of policy and politics, religion and rhetoric, leaving science to concentrate on matters of fact. Nonsense, say the qualitative methodologists. Human experience and its meaningfulness are intrinsically imbued with values; it is thus not possible to separate values out from understandings of human phenomena. To strip human experience of its beliefs, commitments, principles, and passions is to render it other than uniquely social, uniquely human.

Enter Politics

Beyond the arguments about the presence and role of values in social inquiry, the great qualitative-quantitative debate also included strands of political arguments related to the role of social inquiry in society, or, more bluntly, whose interests should be served by social research and evaluation. This is perhaps best illustrated by the character of these political disputes in the evaluation field, as audiences are explicitly named in evaluation studies.

Post-positivist, quantitative evaluation characteristically provides information on the degree to which a given social or educational intervention attained intended outcomes, or the outcomes established by decision makers and program developers. For example, an evaluation of teen pregnancy support and prevention program might be quantitatively evaluated using a quasi-experimental design to assess program impact on the identified outcomes of staying in school, not having a repeat pregnancy, participating in prenatal care, and taking steps toward economic independence. Post-positivist evaluation thus has traditionally served the interests of policy makers and other decision makers charged with the responsibility of setting directions for resource allocation in

the service of social problem solving. In the public arena, these are our elected and appointed officials and staff.

Interpretivist, qualitative evaluation characteristically provides information on the quality and meaningfulness of a program experience, from the multiple perspectives of program staff, participants, and associated family and community members. For example, a qualitative evaluation of the same teen pregnancy support and prevention program might use a mini case study design to illuminate and understand the significance and meaningfulness of these program experiences in the lives of its participants, along possibly emergent dimensions of caring, safety, nurturing, and hope, as well as concrete and material benefits in daily life. This information is most likely to be of value to members of the setting being studied—the staff responsible for program implementation, the local officials responsible for program oversight, and those the program is intended to benefit. These different audiences for evaluation, and social inquiry more broadly, invoked in the great debate a political strand about the role of inquiry in society and about whose questions and interests should be addressed by our work.

Reprise

Those were heady times—exceptionally intellectually demanding, as many of us were challenged to learn a whole host of new concepts and ideas, many of them highly abstract and deep. We read widely, eager for any books or articles that could help us learn as we strove to be among the enlightened. We jammed into conference sessions that featured people in the know. We met in revolt and in protest against revolt, establishing new interest groups and convening small conferences to advance our disparate interests (Eisner and Peshkin, 1990). Heroes and heroines emerged, as did antiheroes and antiheroines. Which was which obviously depended on your point of view. And we argued, sometimes with civility and manners, and sometimes not. One excerpt from an exchange between two presidents of the American Evaluation Society, both evaluation scholars, illustrates the sometimes contentious tenor and contours of this great debate.

> *Whatever it is we have now, it is clearly not working. . . . We need to move beyond cost benefit analysis and objective achievement measure to interpretive realms . . . to begin talking about what our programs mean, what our evaluations tell us, and what they contribute to our understandings as a culture and as a society. We need literally begin to shape . . . the dreams of all of us into realities.* (Lincoln, 1991, p. 6)

> *Groundwater contamination may be simply a construction of some scientist's fertile mind and not real at all, as the Fourth Generation appears to insist, but I would not drink that constructed water myself. If we want to have the maximum likelihood of our results being accepted and used, we will do well to ground them, not in theory and hermeneutics, but in the dependable rigor afforded by our best science and accompanying quantitative analyses.* (Sechrest, 1992, p. 3)

For more in-depth discussions of the substantive issues being debated, see these references, among others: Bernstein (1983); Bryman (1988); Cook and Reichardt (1979); Cronbach (1975); Denzin, Van Maanen, and Manning (1989); Filstead (1970, 1979); Guba (1985, 1990); Hammersley (1992); Hargreaves (1985); Kuhn (1970); Lincoln and Guba (1985); McCarthy (1981); Phillips (1990); Rist (1980); and Smith (1989).

RAPPROCHEMENT AND THE EMERGENCE OF THE IDEA OF MIXING METHODS

Gradually, and grudgingly in some quarters, a truce in the great debate emerged, a truce that legitimized multiple rationales for and multiple ways of practicing social inquiry. Although most social inquirers maintained an allegiance to one particular methodological tradition, most also accepted a plurality of legitimate traditions. And many had substantially deeper understandings of and rationales for their own methodological allegiance or home base as part of their acceptance of other traditions. For example, Nathaniel Gage, an esteemed senior educational researcher, observed that

> *It was finally understood that nothing about objective-quantitative research precluded the description and analysis of classroom processes with interpretive-qualitative methods. Classroom processes need not be described solely in terms of behaviors or actions; they could also be described in terms of meaning-perspectives. (Gage, 1989, p. 7)*

And Lee Cronbach, one of the twentieth century's social science geniuses, observed, regarding evaluation, that "merit lies not in form of inquiry but in relevance of information . . . [and] the evaluator will be wise not to declare allegiance to either a quantitative-scientific-summative methodology or a qualitative-naturalistic-descriptive methodology" (Cronbach & Associates, 1980, p. 7). Typical of this truce was the idea that qualitative methods were good for gathering data on some aspects of human behavior, and quantitative methods were good for gathering data on other aspects of human behavior. That is, the two methodological traditions could function in a complementary fashion, each contributing uniquely to the results (see, for example, Kidder & Fine, 1987; Rossman & Wilson, 1985; Salomon, 1991; A. G. Smith & Louis, 1982; M. L. Smith, 1986). This kind of thinking presaged the turn to the idea of mixed methods social inquiry. In many ways, this turn was one logical outcome of the great debate.

Traditions of Triangulation

The social science community was also receptive to mixed methods ideas because of strong commitments to ideas related to the concept of triangulation in both quantitative and qualitative traditions. From its classic sources, triangulation refers to the intentional use of multiple methods, with offsetting or counteracting biases, in investigations of the same phenomenon to strengthen the validity of inquiry results. The core premise of triangulation is that all methods have inherent biases and limitations—for example, the social desirability known to plague social science surveys and many personal interviews—so that use of only one method to assess a given phenomenon will inevitably yield biased and limited results. However, when two or more methods that have offsetting biases are used to assess a given phenomenon, and the results of these methods converge or corroborate one another, then the validity or credibility of inquiry findings is enhanced. This idea of triangulation is derived from other fields, like surveying and astronomy, where more precise results are achieved by taking measurements from two or more positions.

In quantitative methodological traditions, the concept of triangulation is usually traced first to the multitrait, multimethod (MTMM) matrix conceptualization offered by Campbell and Fiske (1959) for the development and validation of new psychological measures. With this matrix conceptualization, the validity of inferences from a new measure could be convincingly established through a combination of divergent validity (with inferences from measures of different constructs) and convergent validity (with inferences from measures of the same construct), controlling for or separating out shared variance due to similar methods. The actual triangulation label was first used in quantitative traditions by the authors of the highly creative idea of using unobtrusive measurement to support, substantiate, and especially to "cross-validate" more intentional measurement (Webb, Campbell, Schwartz, & Sechrest, 1966).

In qualitative traditions, the triangulation concept is nearly always traced to sociologist Norman Denzin's 1978 text on sociological methods. In this text, Denzin proposed four forms of triangulation—of data sources, methods (specifically interview and observation), investigators, and theories—as ways to overcome limitations of any single data generation perspective or event. Of special focus in this conceptualization of triangulation was the importance of both asking inquiry participants for their interpretations of their experiences and observing the same individuals in action. What people say and what people do are not always the same, and understanding each can inform an understanding of the whole. That is, in qualitative methodological traditions, triangulation was a vehicle to develop a more coherent and comprehensive account or story of the phenomena being studied, as this constituted the interpretive version of validity or credibility.

As noted by Greene and McClintock (1985), the classic triangulation argument requires that the two or more methods be (1) intentionally used to assess the same phenomenon, conceptualized the same way; (2) therefore implemented simultaneously; and (3) also implemented independently, to preserve their counteracting biases. Challenges of triangulation have persisted in the mixed methods literature, as have misunderstandings of what this concept means (Mark & Shotland, 1987b; Mathison, 1988).

And although familiarity with the idea of triangulation may have helped, in some important ways, to support early ideas about mixing methods, the lingering effects of the triangulation concept in the mixed methods conversation have not always been as productive. In particular, triangulation seeks enhanced validity or credibility through convergence and corroboration. There is sometimes a conflation of the very notion of mixed methods social inquiry with the ideas of convergence and corroboration. For example, in my review of early mixed methods evaluation studies, conducted with colleagues Valerie Caracelli and Wendy Graham (Greene, Caracelli, & Graham, 1989), the value of mixing methods for purposes of triangulation was widely promoted, across very different kinds of studies. Triangulation was the dominant stated rationale for using a mix of methods in these studies, even when convergence for purposes of increased validity was not the actual intent of mixing methods in the study's implementation. And a chapter on making inferences in mixed methods inquiry in the *Handbook of Mixed Methods in Social and Behavioral Research* focuses the discussion primarily on triangulated inferences (Erzberger & Kelle, 2003). Muted by this emphasis on convergence and corroboration is the potential value of divergence and dissonance. As noted previously, one important stance in this book is that mixed methods social inquiry can substantially enhance our understanding of social phenomena by generating empirical puzzles—results that do not converge and thereby warrant further study and contemplation (see also Mathison, 1988).

Moreover, the great qualitative-quantitative debate had left simple methodologists with newfound knowledge, if not complete understanding, about the philosophical foundations of their craft. Early on in the mixed methods conversation, questions about the possibility and the sensibility of cross-paradigm triangulation were raised. Greene and McClintock (1985), for example, conducted a self-consciously mixed methods, mixed-paradigm study of the ways in which evaluative data are used in learning and decision making in the educational institution of the Cooperative Extension System. Although each component of the study was conducted relatively independently (meeting one important condition of triangulation), the various sets of qualitative and quantitative results were carefully, even painstakingly integrated at the point of making conclusions and inferences. My colleague Charles McClintock and I then reflectively wondered, what exactly did we do during the process of integration? We knew that different sets of philosophical assumptions had guided the data-gathering processes for each component—because that was an intentional strand of this study—but we did not know what philosophical assumptions had framed and guided the integration process. Our reflections suggested that the integration had occurred in a mostly post-positivist framework, possibly denying the sensibility of cross-paradigm triangulation. Questions about these issues persist today.

Early Mixed Methods Ideas

In addition to renewed scholarly and practical interest in the concept of triangulation, the early mixed methods conversation was imbued with abundant seeds of creative ideas, marking the generative potential of this way of thinking about and conducting social inquiry. Some of these are introduced here as snapshots of the times and as harbingers of the material in the chapters to follow.

First, a handful of instant empirical classics were cited in nearly every early mixed methods article and discussion. These included the Phelan (1987) and Trend (1979) studies featured in Chapter Two. What made a mixed methods empirical study an instant classic was its clear, unequivocal demonstration of insights and inferences that were attained from the mix of methods and that would not have been attained with only one type of method. This notion of an empirical classic is still true today and is well illustrated by Tom Weisner's (2005) edited collection of exciting mixed methods empirical studies from the field of child development. Weisner's collection was assembled for the express purpose of illustrating the innovations of thought possible with intentional and methodologically rigorous mixes of methods in practice. Several of the early classics—perhaps most notably the housing policy study reported by Trend—were additionally augmented by the drama of overt conflict and contestations about data and even investigator credibility. These early mixed methods studies, though not labeled as such, were thus tantalizing enticements of an exciting new approach to social inquiry, with possibly high demands in terms of emotional angst but also potentially important rewards in terms of creative and insightful understandings.

Second, also frequently cited in these early mixed methods conversations were the ideas of sociologist Samuel Sieber (1973), who sought early on to create spaces for the joint use of the two dominant methodological traditions in sociology: surveys and field work. Doren Madey (1982) extended these ideas to the field of program evaluation, with many useful examples. The primary idea promoted by Sieber and later Madey was the value of using one kind of method to help develop the other method, where development included not just actual instrumentation but also sample identification and selection, as well as direction for data analyses. One analysis example offered by Madey was the use of data from qualitative observations and interviews to help construct various indices from the quantitative data sets. The qualitative data served to signal important components of the constructs for which indices were constructed. Just as important, Sieber and Madey also presented these ideas as not only how qualitative methods can strengthen and enhance quantitative methods, but also—in a rare balanced and evenhanded manner—how quantitative methods can strengthen and enhance qualitative methods. An analytic example of this from Madey (drawn from Sieber) was the use of quantitative data to help correct the "holistic fallacy" of many qualitative analysts. The holistic fallacy is the tendency on the part of field observers to perceive all aspects of a situation as congruent. In Madey's evaluation, for example, a site director's positive report about the efficiency of the program's administrative operations at the director's site could have been holistically (mis)understood by the qualitative interviewer, a misunderstanding revealed when the quantitative administrative data indicated site problems with low client participation and high staff turnover. In sum, the work of Sieber and Madey highlighted the contributions possible with a sequential and balanced use of both qualitative and quantitative methods in social inquiry.

Third, other strands of the early discussions of mixing methods in social inquiry featured other possible roles or purposes for mixing methods and other facets of mixed methods design of possible importance. Gretchen Rossman and Bruce Wilson (1985) identified three functions or purposes for a mixed methodology: *corroboration* or convergence, *elaboration* or providing richness and detail (later relabeled *complementarity* by Greene et al., 1989), and *initiation*, which "prompts new interpretations, suggests areas for further exploration, or recasts the entire research question. Initiation brings with it fresh insight and a feeling of the creative leap. . . . Rather than seeking confirmatory evidence, this [initiation] design searches for the provocative" (p. 637 and 633). Extending Rossman and Wilson's ideas from an analysis of evaluation studies that used a mix of methods, Greene et al. (1989) suggested two additional purposes: the use of one method to help develop the other (called *development*, from Sieber & Madey) and the use of different methods for different components of an evaluation study; for example, assessments of program implementation and program outcomes (labeled *expansion*). Bryman (1988) presaged important continuing conversations about the value of multiple perspectives potentially captured by a mix of methods—notably, those of structure and process, outside-researcher and inside-participant, macro and micro levels, and cause and meaning. And John Brewer and Albert Hunter (1989) drew attention to the importance of a mixed methods design that features integration of the different methods throughout the study. (The more recent book by these authors revisits these ideas; Brewer & Hunter, 2005.)

Fourth, Charles Ragin helped to initiate discussions on analysis of mixed data sets. Ragin, an international comparative researcher who studies, for example, the emergence of democracy around the globe, posited two primary approaches to such research: variable-oriented and case-oriented, each with its strengths and limitations. In *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies* (1987), Ragin invented a methodological procedure involving Boolean algebra that enables data from both variable- and case-oriented studies to be combined. Ragin later developed software to support this procedure, foreshadowing contemporary work by Patricia Bazeley on computer software that facilitates iterative exchanges of analyses of quantitative data with SPSS and qualitative data with NVivo (Bazeley, 2003, 2006).

The final snapshot of these early mixed methods discussions returns to the "paradigm issue" in mixed methods inquiry. Rossman and Wilson (1985) offered three stances on the questions of whether and how philosophical paradigms can be mixed in mixed methods research: (1) the purists who say "Absolutely not," (2) the pragmatists who say "Of course, what's the problem here?" and (3) the middle-ground situationalists who say "Maybe, especially if we reframe the notion of philosophical paradigm." These ideas are engaged in some detail in Chapter Five. Related to these issues, Reichardt and Cook (1979) presented a catalytic set of ideas regarding the nature of philosophical paradigms and their relationships to practice. They challenged both the inviolability of philosophical paradigms and their directive role in social research. This set of issues is taken up in the next chapter as a prelude to Chapter Five's direct engagement with the paradigm issue.

Additional influential contributors to these early conversations about mixed methods social inquiry include Mark and Shotland (1987a), Fielding and Fielding (1986), Smith and Louis (1982), and M.L. Smith (1986), some of which were anchored in more general seminal work such as Mary Kennedy's (1979) paper on generalizing from the case study and Donald Campbell's classics on qualitative knowing in action research (1978) and on degrees of freedom in the case study (1979).

FOR EXAMPLE

An illustration of an exciting mixed methods study being conducted at this time of early rapprochement is the dissertation on understanding hunger conducted by Kathy Radimer (1990). Radimer positioned her study in the midst of a highly politicized debate about the existence of hunger in America during the 1980s. She argued that the debate was rooted in "lack of agreement as to what constitutes hunger and what indicators are considered to signify it" (p. 50). She then argued that the meanings of hunger should be grounded in the contextualized experiences of those who "go hungry," and it is these meanings that should be translated into measurable indicators for national survey purposes.

Radimer's dissertation then proceeded to enact this argument through a two-part study. In part one, she conducted in-depth interviews with thirty-two women identified through community-based organizations serving low-income families. Extensive and iterative analysis of these qualitative data yielded a complex portrait of the meanings of hunger for this sample. This portrait included three major dimensions of the experience of being hungry, each with multiple facets and characteristics: problems with the quantity of food intake, problems with the quality of food intake, and problems with food supply for the household overall. A fourth dimension or closely related construct described by the interviewees was food anxiety. The interview data also yielded rich information on how these women manage food insecurity and the threat of going hungry, which Radimer included, along with still other significant dimensions and events, in her overall conceptual framework for hunger.

In part two of her dissertation, Radimer developed survey items based on the conceptual framework for hunger that emerged from the qualitative data and pilot tested these items on a sample of women with a broader range of hunger experiences, from none to substantial. Extensive psychometric analyses were used to assess the congruence of the pilot test results with the conceptual framework and to assess the reliability and validity of intended inferences about hunger from the survey. Survey results were also used to make modest revisions in the conceptual framework, which was then used to recommend a small set of hunger indicators for use on national surveys assessing the state of hunger in the nation.

Radimer's study is exemplary in its in-depth engagement with the complexities of a major public health issue like hunger. Her study well illustrates the conceptual and practical value of a mixed methods approach and especially one with open respect for the legitimacy of multiple ways of knowing and understanding.

BUT TROUBLED WATERS REMAINED

These snapshots from the early conversations about mixing methods in social inquiry are intended to convey some of the exciting sense of potential that followed the proclaimed rapprochement of the great qualitative-quantitative debate. But murmurs of challenge remained, some of which continue as critical strands of the mixed methods conversation today. These include the following:

- Some inquirers with allegiance to qualitative traditions were concerned that their particular way of knowing was actually being silenced by the interest in mixing methods, which they perceived as a thinly disguised advocacy for continued post-positivist and quantitative dominance.
- Some participants wondered what all the fuss was about, as some social inquirers, like program evaluators, had been fruitfully mixing methods for quite some time (Datta, 1994).
- Other participants in the early conversations were concerned that the rush to rapprochement and partnership (Reichardt and Rallis, 1994) left many important issues and concerns by the wayside. Philosophically, for example, the long-standing challenges of incommensurability (Howe, 1988, 2003; J. K. Smith, 1989) were not sufficiently well engaged at that time.

These and other issues will be pursued as this conversation moves on from history. The next two chapters specifically engage the paradigm issue in mixed method social inquiry.

CHAPTER 4

CONTESTED SPACES: PARADIGMS AND PRACTICE IN MIXED METHODS SOCIAL INQUIRY

THE LEGACIES of the qualitative-quantitative debate in applied social science persist today, as do some of the tensions, especially regarding the nature and role of philosophical assumptions and stances in social inquiry practice. Back from the quick trip to the past, in this chapter the traveler will venture out into the abstract and roughly textured landscape of philosophical paradigms and mental models. The traveler will explore a number of different routes through this landscape *to* the domain of social inquiry practice and will begin to map the route most congruent with his or her own travel modes, speeds, and destinations. The traveler will further