

# Toward a General and Unified View of Educational Research and Educational Evaluation

## Bridging Philosophy and Methodology

Rocco J. Perla

*George W. Merck Fellow*

*Institute for Healthcare Improvement*

James Carifio

*University of Massachusetts—Lowell*

**ABSTRACT:** This paper addresses an important problem that may really be a pseudoproblem perpetuated by a current artificial or even political need to isolate researchers, evaluators, and theorists and put them in conceptual-methodological boxes. In this paper, we argue that scholars of all kinds should focus on the nouns (i.e., research, evaluation, theory) versus the adjectives (i.e., quantitative, qualitative, or mixed methods) used to qualify these nouns and core constructs. In a word, researchers, evaluators, measurement professionals, and theorists of all persuasions should strive to be neither qualitative nor quantitative, as that is like striving to be a single leg in a triangle. Following Hacking (1965), our view is that all forms of research methodology and evaluation methodology represent a “patchwork quilt” that is essentially “one fabric” with two connected but often unacknowledged sides to the quilt and that all methodology, analysis, and evaluations are essentially “qual-quantification” and “quant-qualification” and usually both at the same time. This general and unified or “one fabric” view of these debates and issues is discussed from a historical-philosophical perspective, and the need for a more general and inclusive evaluation and research epistemology is elucidated. Having outlined an epistemic framework and argument for a more comprehensive and complete view and model of evaluation and research that transcends a specific research paradigm, the paper concludes by outlining eight general key points that most (if not all) researchers and evaluators should consider.

**KEYWORDS:** *education; epistemology; measurement; qualitative research; quantitative research; research and evaluation; meta-theory; mixed methods research*

When it comes to contentious subject matter in academia, one would be hard pressed to find a more sensitive, ideological, and volatile topic today than research and evaluation methodology, and particularly the ongoing debate between qualitative and quantitative research and evaluation techniques (Coryn,

2007; Pawson & Tilley, 2007). As Glesne (1999) points out, “The research [or evaluation] methods you choose say something about your views on what qualifies as valuable knowledge and your perspective on the nature of reality or *ontology*” (p. 4). As the distinctions between these two (currently competing) research traditions

becomes sharper and more distinct, the greater the lacunae between not just the methodological *modus operandi*, but also their associated epistemological, ontological, and philosophical views and assumptions. The “achieved clarities” of these different research methods has paradoxically led to a diametrically opposite view of what nuanced and sophisticated (or “good”) research and evaluation is and should aspire to be. This reductionist approach to methodology is not only divisive, but tends to trivialize many of the important and central unifying themes and principles of high quality research and evaluation programs (Coryn, 2007). Hence, there is a need for a more general and unified view and theory of research and evaluation that transcends a particular method and focuses on research as a way of thinking and specific habits of mind (i.e., an epistemology). The ideas developed in this paper have begun to emerge in some of the seminal works done on mixed research and measurement (Johnson & Onwuegbuzie, 2004), some of the redefinitions and reconceptualizations of experiments (Sloane, 2008), Pawson and Tilley’s (2007) realistic evaluation, Stufflebeam and Shinkfield’s (2007) generic principles underlying sound evaluation procedures and techniques, and Kleinig’s (1982) work on qualitative heuristics that attempts to reintroduce the qualities of exploration and discovery into the research and evaluation process and the concept of the qualitative experiment. As much of the research/evaluation methodology debate has stemmed from the war between the qualitative and quantitative camps, we focus on these two camps as the starting point in our analysis.

The debates and purported differences between the camps of qualitative and quantitative methodologists are well described (e.g., Denzin & Lincoln, 2005), while a more middle-of-the-road approach—the mixed-method approach—has been advanced in recent times and is gaining in popularity (Johnson & Onwuegbuzie, 2004). Each of these three

different traditions claims philosophical, epistemological, and ontological autonomy. Qualitative methods are often characterized (by some) as more reflective, interpretive, subjective, empathetic, contextual, social, and theory-less (or theory-forming), while quantitative methods are described (by others) as more calculated, precise, rational, theory-driven, predictive, generalizable, objective, testable, empirical, and positivist-like (Glesne, 1999; Kerlinger & Lee, 2000). The mixed-method proponents argue that their program combines different aspects of each of these approaches to obtain the best of both possible worlds in a pragmatist framework. However, one of us has long argued that the “patchwork quilt” (Hacking, 1965) that is both methodology and analysis is essentially “one fabric” with perhaps two connected, but often unacknowledged sides; and all methodology and analysis is essentially “qual-quantification” and “quant-qualification” and usually both at the same time. The other of us arrived at this same view through philosophical analyses and perspectives. This general and unified “one-fabric-even-if-a-patchwork-quilt” view of these debates and issues will be pursued in more detail at the end of this paper.

The traditional differences cited above between qualitative and quantitative methodologies and commitments only provide a source of unequivocal distinction for some advocates. But without such sharp (if somewhat artificial) divisions, the lines between these traditions would be blurred and their identities lost. Assuming that what are called the qualitative and quantitative traditions are interwoven—and that they compliment and extend one another (versus being exclusive) in a dynamic and nonlinear manner—would likely lead to some initial confusion and chaos (similar to what Kuhn (1962/1996) referred to as a phase of “extraordinary science”) followed by a better view of what sound research is and should be. This imbrication of methodology, anticipated in numerous philosophical and

epistemological writings (see below), would also make the mixed-method approach and designation moot.

The term and concept “mixed method” as used today (see Johnson & Onwuegbuzie, 2004) involves using *either* a qualitative method or perspective *or* a quantitative method or perspective in response to the requirements of the research at different points in time. However, this methodological perspective leads to the false implication and entailment that nothing that is characterized as quantitative is qualitative and vice versa. The problem with this implication is that much of the foundational theory and insight that defines modern statistics and psychometrics is very consistent with what is called qualitative methods or “non-experimental wisdom” (Campbell & Stanley 1963, p. 3), whereas many of the concepts and terms used in qualitative research are derived from psychometric theory and theories of statistical inferencing. For example, qualitative researchers often refer to the importance of “triangulation” among different sources of information as a verification procedure of research findings (Creswell, 1994). This verification process is consistent in principle with the convergent and discriminant design used by many quantitative researchers (see Campbell & Fiske, 1959, for details) to assess construct (i.e., theoretical) validity. There are many more examples of conceptual and linguistic similarities and equivalences that can be recovered from qualitative and quantitative research traditions. These types of issues may contribute to and partly explain some of the observed barriers to the integration of qualitative and quantitative methods in research programs (Bryman, 2007).

As some have recognized, strict hierarchies of method where the randomized controlled trial (RCT) is king regardless of context is not the most productive evaluative framework (Brass, Nunez-Neto, & Williams, 2006). However, the process of acknowledging the value of different research methods has the

unanticipated consequence of creating more division through legitimization; the fact of the matter is that there is good and bad research and program evaluation regardless of what someone calls it. Without recognizing this, the world of research and program evaluation has no set of first principles.

Instead of focusing on similarities and eliminating redundancies of methodology, research and evaluation communities—and particularly some educational and social science communities—have been more focused on making claims of distinction and separation versus identifying general principles of nuanced and sophisticated or sound and valid research. In a historical context (ca 1930s), much of the discontent with experimental models in education and social science stems from the grandiose claims of early experimentalists and their failure to produce the types of results anticipated, where experiment is defined as “that portion of research in which variables are manipulated and their effects upon other variables observed” (Campbell & Stanley, 1963, p. 1). Indeed, as Campbell and Stanley point out, some people *well trained* in the experimental tradition have periodically defected “from experimentation to essay writing” (p. 2) based in part on a disillusionment of success associated with early experimental models and the unrealistic expectation that a scientific and quantitative approach would ameliorate all major problems in education, educational research, and assessment.

There is no question that there exist different ways to approach and execute different parts of a research study or evaluation from conceptualization to falsification and all points in between. But we should teach students as well as future researchers and evaluators to match the best methodological and evaluative tools to the particular problems at hand—not eliminate one-half (or more) of the toolbox. As Pawson and Tilley (2007) point out in the context of a realistic research design, there is “no one standard ‘formula’, other than the base

strategy of producing a clear theory of program mechanisms, contexts and outcomes, and then using them to design the appropriate empirical measures and comparisons” (p. xv). Further, the tools we use are not unrelated at the epistemological level, as most if not all of the quantitative methods available can be described in qualitative-like terminology to varying degrees and vice versa. Indeed, any view of quantitative research and evaluation as just “number-based” is a poor, naïve, and simply incorrect view of the intent and form of quantitative research, theory, and practice. Similarly, the view that qualitative research and evaluation does not and cannot anticipate rigorous statistical or psychometric testing (leading to greater control over variables and reducing sources of invalidity and error) is unfortunate and an unnecessary and often self-induced limitation on the parts of many qualitative methodologists. In this sense, then, we argue that the divisions between quantitative, qualitative, and mixed-methods research and evaluation is, and has been, artificial and somewhat akin to what Ryle (1949) classically defined as a category mistake—a cognitive or reasoning error that involves placement of one concept or concepts in an incorrect category or the description of a concept using fallacious, inappropriate, or misaligned criteria. And, as Ryle so forcefully argued, some category mistakes appear subtle, but can have devastating consequences and implications (such as the Cartesian separation of mind and body).

The basic thesis advanced here is that *philosophy, epistemology, and history matter* and are foundational to any sophisticated view and program of measurement, evaluation, and research. It is our view that many practitioners and students may know how to operationalize a research and evaluation method—particularly the types in which they have been inculcated—but most are unaware of the history, richness, and complexity associated with the method’s underlying epistemology. Understanding how

the methods one uses come together to form a *meaning-making system* is the very bloodline of quality research and evaluation, and it is also the essence of epistemology. Without some understanding and appreciation for the epistemic and philosophic issues that surround a particular methodology or paradigm, it becomes easy to create superficial distinctions of method that upon careful examination migrate to one end of the same spectrum (versus creating a different way of doing research and evaluation or doing some “quantum tunneling” to the other side of the quilt).

The purpose of this commentary is to identify general characteristics and considerations of “good” (nuanced, sophisticated, and valid) research and evaluation that transcend (or at least “fuzzify”) the assumed boundaries of the different paradigms extant today. It is neither practical nor necessary for our argument here to address each epistemic issue or referent that is central to research methodology, measurement, and evaluation. Instead, we look to begin a discussion of how history, philosophy, and epistemology inform and broaden one’s perspective on research methodology and evaluative processes at a very fundamental level. Indeed, we are not advocating any one particular form of research or evaluation, merely that the methods one uses in a particular context is part of a larger meaning-making system that should, in theory, lend itself to different methods based on where the inquiry ends up. It should be noted at this point that “quantitative methodologists” are often not on the solid footing they perceive, while “qualitative methodologists” are often closer to the truth of the matter. Some may believe the basic thesis developed here relative to disputes of method are well-established and therefore unnecessary or naïve or even hallowed ground. Although it is true that these issues are venerable in character, dating back to Descartes’ classic discourse on the subject,<sup>1</sup> the sentiments we are most concerned about in a research and

evaluative context are still alive and well to say the least. By analyzing a sample of these key issues using classic and primary philosophical referents, we hope to encourage a deeper level of thought and reflection on the subject.

We begin our analysis briefly with a classic example of how epistemology informs methodology and the blurring of the supposed boundaries between the “exact” and “inexact” sciences by reviewing the genesis of the Delphi methodology developed during the 1950s and 1960s by the RAND cooperation. We then address an important philosophical issue that is central to understanding *comprehensive meaning-making systems*: the contexts of justification and discovery. This section is followed by a discussion of logical positivism and its association (and often ignored relationship) with more qualitative-like research activities and methods and the context of discovery. This section highlights four critical reasoning errors that antipositivists often make in their rejection of the positivist philosophy and epistemology that contribute to the present-day isolationism of qualitative and quantitative methods and research. We conclude by identifying eight key points and issues related to epistemologically, empirically, and theoretically sound research and evaluation that encompasses what has traditionally been called qualitative and quantitative methodology.

## The Delphi Epistemic Foundation

In a paper that laid the epistemic foundation for what later became known as the Delphi method, Helmer and Rescher (1960) argue that the distinction between the exact and inexact sciences has been confused, similar to a Rylean category mistake as mentioned above. In fact, Helmer and Rescher state the following:

It is a fiction of longstanding that there are two classes of sciences . . . . This widely prevalent attitude seems to us fundamentally mistaken; for it finds a difference in principle where there is only one of degree, and it imputes to the so-

called exact sciences a procedural rigor that is rarely present in fact. (p. 1)

The classical notion of the exact sciences, as noted by Helmer and Rescher, involves a reasoning process where terms are exactly defined and the reasoning occurs by a formal logico-mathematical derivation of the hypothesis based on the available evidence. However, the central point to Helmer and Rescher’s (1960) view is that “while precise predictions are indeed to be preferred to vague ones, a discipline that provides predictions of a less precise character, but makes them correctly and in a systematic and reasoned way, must be classified as a science” (p. 1). Helmer and Rescher clearly understood that their newly invented Delphi technique was a methodology that was less precise in character, but that it was a methodology that made predictions correctly, systematically, and in a reasoned manner. The Delphi method, therefore, was scientific by Helmer and Rescher’s definition and clearly within the scientific tradition, although it is typically called a qualitative methodology today. This example needs to be well-noted, as does the Rylean reclassification of the Delphi method that has occurred since its original development—assiduously outlined and justified in epistemic terms—which seems to have gotten lost in the historical haze of various disciplines including education.

Given Helmer and Rescher’s views above, there are two important points we need to make here relative to the topic of this paper. First, as noted almost a half century ago by Helmer and Rescher (1960), the differences between the exact and inexact sciences are indeed differences of degree rather than type if one adopts a more general and realistic view of science (albeit one that continues to include prediction and a systematic approach to explore phenomena of all kinds). Second, all researchers need to understand the epistemic foundation and development of their research choices and programs. Indeed, if research and evaluation are viewed as meaning-making processes, then one

cannot appreciate the significance, implications, and limitations of these processes and activities without knowing something of the epistemic framework from which they come and work through. This lack of epistemic understanding of research processes is, in our opinion, one of the reasons why armchair speculation and blind (shotgun) empiricism—whether quantitative or qualitative—are so problematic today in many forms of healthcare, social science, and educational research. This epistemic theme and focus is continued in the next section of the paper that addresses two important—and venerable—philosophical contexts rarely discussed in the setting of contemporary research and evaluation programs.

## Context of Discovery *and* Context of Justification

Both the context of discovery and the context of justification have a long history in philosophy, particularly as epistemological concepts, and one could not exist without the backdrop and contrast of the other (Reichenbach, 1938). It is important to note that the term epistemology, as it is used by philosophers in a traditional sense, most often refers to the context of justification or the methods of rational reconstruction of knowledge that are *prescriptive* in that they dictate how *we should think*, while the context of discovery was originally seen as a psychological concept and issue that is *descriptive* of how *we actually do think*, rarely with the two (intentionally) crossing intellectual paths (Reichenbach, 1938). The goal of epistemology, in a philosophical sense, is to

construct thinking processes in a way in which they ought to occur if they are to be ranged in a consistent system; or to construct justifiable sets of operations which can be intercalated between the starting point and the issue of thought processes, replacing the real intermediate links. (Reichenbach, 1938, p.5)

Epistemology, therefore, “considers a logical substitute rather than real processes” (Reichenbach, 1938, p. 5). The logical substitute is the set of procedures and methods determined a priori that create the mold and standard for rational reconstructions and the expectations of logical (optimal) human reasoning. The real processes are the actual observations of how humans reason that often deviate or vary from the logical expectations.

There is a natural and important correspondence between the contexts of discovery and justification and guiding statistical and psychometric principles. In statistical terms, the distinction between *what is observed* and *what is experienced/expected* represents a critical distinction and ratio that is the foundation for the analysis of variance (ANOVA), regression analysis, and correlation—fundamental statistical as well as psychometric concepts. Whereas traditional philosophers and logicians focus exclusively on *what is expected* (in an idealistic sense) or the context of justification, qualitative methodologists tend to focus exclusively on *what is observed* or the context of discovery.

The problem with focusing exclusively on logical expectations *or* our observations is that it dissolves and negates the underlying key distinction and ratio and eliminates the systems of checks and balances we use to gauge the successes and failings of the theories that guide our research and evaluation programs. The very nature of theory (see Suppe, 1974; Lakatos, 1970) is that it relies on both what we expect to observe (based on the explanative components of the theory) and what we actually observe (based on perception and experience and the methodological components of the theory). Modifying our (often theory-driven) expectations based on our (somewhat controlled) experiences and observations is and has been the basis for knowledge development in modern times and especially in science (Suppe, 1974; Hanson, 1958). This process and cycle of knowledge (i.e., theory) development,

moreover, is not just restricted to the hard sciences, but has now become an expectation of the social sciences (Aneshensel, 2002), educational research (Shavelson & Towne, 2002), and modern evaluation programs (Pawson & Tilley, 2007). This observational fact, however, does not and should not vitiate the fact that qualitative-like research can contribute substantively to theory and knowledge development, albeit in a different yet conceptually rigorous manner. For example, the emphasis might be on creating a framework for creating and subsequently testing a new view, model, conjecture, or idea (see Perla, 2006). Suffice it to say, there can be little to no growth to our research and evaluation programs and theories if they do not include, recognize or address the cyclic relationship between the contexts of discovery and justification and their respective elements and subjudgments.

One of the admitted limitations of early analytical philosophers and logicians and their formal systems and models of knowing (especially the logical positivist program) is that despite the fact they recognized the important role and impact of the extra-scientific and nonexperimental wisdom on the knowledge development process, they excluded it from their analysis. Early logical positivists were content to accept the things that “could not be observed *and* measured” and to characterize them as meaningless and simply eliminate them from their formal calculi and conceptual maps, even in a fuzzy and probabilistic sense. In fact, the early positivist assumption that a complete (100%) correspondence could exist between what we observe and the symbolic language used to represent these observations was the motivation for Reichenbach’s (1938) criticism of logical positivism and his model of probabilistic empiricism. That model addressed the inescapable fact that the correspondence between symbolic language and observation is never 100 percent or exact and is always a partial correspondence, with the key estimate being how partial or incomplete.

In developing his thesis in *Experience and Prediction*, Reichenbach (1938), who coined the expressions “context of discovery” and “context of justification,” emphasized two important points. The first point was that the context of justification or rational reconstruction was never perfect and always subject to vagaries of human language and inexactitudes. The second point was that scientific explorations and research often begin with certain choices (volitions), conventions, and heuristics that are not governed by logic as much as by axioms of choice. As Reichenbach points out:

Scientific method is not, in every step of its procedure, directed by the principle of validity; there are other steps that have the character of volitional decisions. It is this distinction which we must emphasize at the very beginning of epistemological investigations. That the idea of truth, or validity, has a directive influence in scientific thinking is obvious and has at all times been noticed by epistemologists. That there are certain elements of knowledge, however, which are not governed by the idea of truth, but which are due to volitional resolutions, and though highly influencing the makeup of the whole system of knowledge, do not touch its truth character, is less known to philosophical investigators. (p. 9)

These “volitional resolutions” were not as much unknown to philosophical investigators of that time as Reichenbach seems to suggest. They were known, but were virtually impossible to understand due to their admittedly complex structure and a dearth of conceptual and methodological precedents for dealing with them in any meaningful way *at the time*.

A philosophical awareness that certain elements of knowledge are not guided by the notion of truth and a formal calculus does predate Reichenbach’s book, most notably in the work of Ludwig Fleck (1935/1979), a Polish-born physician and epistemologist. Consider the historical findings of Fleck in his classic book, *Genesis and Development of a Scientific Fact*, that is one of the most thoughtful and

comprehensive accounts of the development of the modern concept of syphilis from its mystical character and origins to its etiological and pathological (disease-causing) character:

Many very solidly established scientific facts are undeniably linked in their development, to prescientific, somewhat hazy, related proto-ideas or pre-ideas, even though such links cannot be substantiated. (p. 23)

...a proto-idea must not be construed as a “freak of nature.” Proto-ideas must be regarded as developmental rudiments of modern theories and as originating from a socio-cognitive foundation . . . . The value of such a pre-idea resides neither in its inner logic nor in its “objective” content as such, but solely in the heuristic significance which it has in the natural tendency of development. And there is no doubt that a fact develops step by step from this hazy proto-idea, which is neither right nor wrong. (p. 25)

The acquisition of physical and psychological skills, the amassing of a certain number of observations and experiments, the ability to mold concepts, however, introduce all kinds of factors that cannot be regulated by formal logic. (p. 10)

Consonant with these views, the logical positivist Phillip Frank (1949) later stated,

If we want to evaluate precisely and critically how firmly this philosophy [of science] is anchored in the ground of science, we must not ignore the extra-scientific factors, but must analyze carefully the social, ethical and religious influences. Every satisfactory philosophy of science has to combine logic of science with sociology of science. (p. i)

Later Quine (1961) argued that rejecting the “two dogmas of empiricism” (i.e., the cleavage between analytic and synthetic truths and the notion that all meaningful statements can be reduced to a logical statement about direct experience) leads to “a blurring of the supposed boundary between speculative metaphysics and natural science” (p. 20). In commenting on what

he referred to as the pseudoscientific (metaphysical/untestable) theories of Marx, Freud, and Adler, Karl Popper noted that he “realized that such myths may be developed, and become testable” and that “historically speaking all—or very nearly all—scientific theories originate from myths, and that a myth may contain important anticipations of scientific theories” (as cited in Schick, 2000, p. 12). Although Popper was not concerned with the nature of this transition per se, he clearly recognized the fundamental importance of what Fleck (1935/1979) describes as a proto-idea and the interdependence of what are now referred to as qualitative and quantitative forms of knowing.

As these brief examples demonstrate, in any comprehensive modern research and evaluation program, both the contexts of discovery and justification should be addressed, considered, partly integrated, and anticipated even if not operationalized. Knowing that many scientific and systematic ideas develop over time from highly amorphous, qualitative-like concepts and with the advantages in technology that allow us to approach metaphysical questions of the mind previously described as meaningless (Pinker, 1997), dismissing difficult questions of the nonlogical variety is no longer a viable or intellectually responsible option. That is, we can no longer be purely philosophical or quantitative or rational in a vacuum. Perhaps the positivists were wrong in relegating all metaphysical speculations to the category of meaningless questions; perhaps the available methods for systematically and logically dealing with metaphysical phenomena were meaningless, or at least so underdeveloped or unrecognized as to be *virtually* meaningless. Cronbach (as cited in Lincoln & Guba, 1985) expresses a similar concern with the positivist-like epistemology that focuses exclusively on the verification principle and neglects the context of discovery:

“Design of experiments” has been a standard element in training for social scientists. This



training has concentrated of formal tests of hypotheses—confirmatory studies—despite the fact that R. A. Fisher, the prime theorist of experimental design, demonstrated over and over again in his agricultural investigations that effective inquiry works back and forth between the heuristic and confirmatory. *But since he could offer a formal theory only for the confirmatory studies, that part came to be taken for the whole.* (p. 25)

The point to be made here is that some forms of qualitative-like research may be able to tease out and identify difficult and complex issues and ideas, which in some cases may be the first step in conceptualizing a method to verify the significance and magnitude of findings and bring the research to the next level. This appears to be the case with Gell-Mann's early conceptualization of subatomic particles, Kekule's well known dreamlike vision of the three dimensional structure of benzene, Ebbinghaus's ground-breaking work on human learning and memory, and Darwin's theory of natural selection and evolution—transformational ideas that were severely ridiculed by colleagues at first, but later experimentally validated and extended.

In pursuing these lines of inquiry, particularly in their more scholarly and historical forms, and trying to maximize the benefits and reduce the risks of research and evaluation, a more general epistemology is needed. This more general epistemology is what this article advocates in nascent form. Recognizing the important advances in the cognitive, computer, and neurosciences, many modern philosophers and most cognitivists have adopted a more general, dynamic, and *interdisciplinary* epistemology that associates the nature of knowledge with both discovery and justification (e.g., Thagard, 1988). What does this broader and more interdisciplinary epistemology translate to from a research methods perspective? At a minimum, it means that the qualitative-like researcher should at least know how to develop and anticipate a plan and model to justify, test, and falsify their findings, views, and theses, while the quantitative-like researcher

should appreciate and understand the types of questions, speculations, and implications associated with their findings and views that may not be immediately amenable to falsification procedures. In other words, we need nuanced and sophisticated researchers, not qualitative or quantitative researchers. Or as Phillips (2005) has so eloquently put it, we need professionals and practitioners who are experts in the conduct and evaluation of each and each type of case. In a grammatical sense, we need to focus on the nouns (research and evaluation) versus the adjectives (quantitative and/or qualitative).

Today, the context of discovery is recognized as the reflective, interpretive, subjective, empathetic, contextual, social, speculative, fuzzy, creative, dialectical, and generative processes used explicitly or implicitly to develop theoretical insights, views, or models. The context of discovery is, therefore, consistent with qualitative research. Conversely, the context of justification is associated with the formal, logical, and more objective methods used to test and validate theoretical insights, views, or models and is, therefore, consistent with the context of justification (or rational reconstruction). The point is that both contexts are important and necessary to understand knowledge systems and knowledge development meaningfully and with the breadth of understanding that should be common to all researchers, which is our main argument. Certain tests and validation procedures may clarify our understanding of a subject or lead to other questions that cannot be answered at the present time. Understanding that the "error term" in our analyses and models included all of the casual and contextual factors we have not included and are unaccounted for in our results, and acting accordingly in interpreting our results, incorporates the qualitative context of discovery (and the need for more sophisticated explorations) explicitly and firmly in our work, interpretation of results, claims and recommendations. Suffice it to say that the

dynamic interplay between the context of discovery and the context of justification defines research agendas and much of our daily experiences and predictions. There is perhaps no better and important historical or epistemological referent or unifying epistemic theme to qualitative and quantitative research today than the context of discovery and justification.

*But this insight has consequences. This insight means that we have to reconceptualize research methodology and evaluation.*

But why do we need to reconceptualize research methods and evaluation using the context of discovery and justification? Why not just broaden our view of qualitative and quantitative methods? The problem is that many of the central (philosophical) arguments used to define the division between qualitative and quantitative methods have traditionally been so misinformed and distorted in their attempts to associate themselves with a particular research method that they serve largely to misinform and create far more problems than they might clarify, let alone solve. One prime example of this distortion is the trivialization, caricature, and misrepresentation of the logical positivist program by some qualitative researchers. To the extent that qualitative research is associated with a distorted view of positivism as described below is enough to render the entire program of qualitative research virtually meaningless and without any theoretical or conceptual foundation. The next section of this paper, therefore, addresses this issue in the form of four main problems and inconsistencies that routinely plague qualitative researchers with antipositivist views.

## Positivism and the Context of Discovery

It is important to point out that the cyclic relationship between fashionable research methods in education is not all that different

from the cyclic transitions between the philosophical movement known as logical positivism and its focus on formal logic, validation procedures, and the context of justification (e.g., Carnap, 1937; Wittgenstein, 1926) and postpositivism and its focus on the psychology of research, generative claims, and assertions and the context of discovery (e.g., Fleck, 1935; Kuhn, 1970). Indeed, temporal and conceptual similarities exist between the formalisms of logical positivism (ca 1920s and 1930s) and its primary aim to “avoid the traditional ambiguity and obscurity of philosophy” and to “bring about the closest possible *rapprochement* between philosophy and science” (Frank, 1949, p. 1) and the formalisms of early experimentally minded educationalists (e.g., McCall, 1923) and their response to the intuitive, informal, and capricious models of educational research of the time. But the formal systems in education and philosophy developed during the early portion of the twentieth century built largely on the context of justification did not hold up to what were unrealistic expectations of success despite the fact that these formal systems were revolutionary as well as ambitious, leaving staunch supporters of psychometrics and positivism to question seriously the value and foundation of their respective programs (see Campbell & Stanley, 1963). The void created by the so-called failings of positivism and psychometrics was filled partly by a postpositivist philosophy of science that was more subjective and relativistic (e.g., Kuhn, 1962) and more qualitative models of educational research similar to those being advanced today.

It is instructive to note that Kuhn’s brand of philosophy of science (often described as antipositivist) has served as a primary and foundational referent for contemporary social scientists (Matthews, 2004a) and as a philosophical and epistemological justification for qualitative research methods, thereby demonstrating a strong relationship between philosophy of science and research

methodology in the social sciences and education. However, many social scientists, educators, and qualitative researchers who use postpositivism (and Kuhn) as a foundational epistemic referent seem to make four critical errors. First, they tend to be selective in that they focus on the effect and implications of Kuhn's "extraordinary" and "revolutionary" science that is more aligned with the context of discovery than on Kuhn's "normal science"—the latter of which is a far more time-dominant phenomena defined by the context of justification and the process of formal logic and axiomatic systems. Any researcher who ignores the impact (and necessary contrast) of normal science to revolutionary science is living in a time warp and ignoring 99.9 percent of Kuhn's data and the data of most historians of science who clearly document that scientific revolutions and "extraordinary science" are extremely rare events compared to normal (everyday, puzzle-solving) science.

Second, many researchers—not just social science or educational researchers—fail to acknowledge or intellectualize the fact that the early "formalist programs" in philosophy and in educational measurement involved a messy, less formal and highly social (qualitative-like) discovery phase. Even the early logical positivists (i.e., the Vienna Circle) disagreed and debated fundamentally important questions and issues. As one of the Circle's members points out in a reflection:

What unites its members is . . . not so much definitive views or dogmas as definite tendencies and endeavors. An evidence of this is the often considerable divergence and lively discussion between its members and the amendments in the fundamental views that have occurred several times in the course of its development. On the other hand, the constant change of opinion has led to an increasing convergence toward certain basic principles that have gradually taken shape and that now form the common basis for the further discussion of still unsettled questions. (Joergensen, 1951, p. 1)

This view of the positivist program has been expressed by other members of the Vienna Circle (e.g., Frank, 1949; Kraft, 1953) and reflects what today is often referred to as qualitative networking and exploratory analysis—a far cry from the dogmatic, rigid, axiomatic and "evil" positivist straw man reconstructed in the social science and educational literature (Matthews, 2004b). Also, Joergensen's comment above clearly suggests an exploratory and more qualitative phase and dimension of the Circle (see second sentence) and movement toward a well-articulated and more quantitative view (see third sentence). In this sense, the positivists were both qualitative and quantitative. Because the positivists' views and main thesis did in fact "converge toward certain basic principles" that could be tested in an extremely clear and rigorous fashion (like the research of Ebbinghaus), their early phase of discovery, exploration, and disagreement is completely ignored—perhaps out of intellectual sloppiness, the straw man tactic, or little familiarity with the intellectual and scholarly history of these issues and questions. Today, much of the positivist epistemology has been falsified due to the unrealistic goal of guaranteed knowledge (dating back to Reichenbach in 1938) and the voluminous *experimental* literature in the cognitive sciences and related fields such as linguistics and mathematical psychology that demonstrates that humans are not reliable Bayesean probability agents (e.g., Tyversky & Kahneman, 1981). But the falsification of some aspects of positivism (and the retention of other aspects) was only possible because these aspects could be tested. Note clearly the word **tested**—the irreducible and inescapable bottom line of an inquiry, claim, argument, or contention (Campbell & Stanley, 1963). In fact, much modern work in the area of scientific discovery has shown that a focus on the context of justification (i.e., rigorous testing and explanation) can both initiate and drive discovery (i.e., the context of discovery) and that these two contexts are not

linearly sequential as erroneously portrayed by many, but rather are bidirectional when employed by researchers who are steeped in both contexts, methodologies, and traditions (see Gigerenzer, 1991).

Herein lies a fundamental problem for any “anti-quantitative” researcher who uses logical positivism as a negative reference to support their view and research method; namely, the third critical error. An anti-quantitative/anti-experimental researcher who is antipositivist has not only accepted the falsification of positivism presumably, but their research ideology and epistemology is actually based on the falsification principle, yet their own practice is in many ways not falsifiable! *Accepting falsification of another view while one’s own view is not falsifiable or using a set of criteria to evaluate and reject an alleged rival view and not to use the same criteria on one’s own view—is a logical contradiction of immense proportion and indefensible intellectually.* It is ironical to observe here that this type of antipositivist view could not, in principle, exist in the absence of the central tenet of positivism: falsification.

The fourth problem that some antipositivist qualitative researchers fail to recognize is that they share the same fundamental view of the nature of observations made in the natural world espoused by positivists. For example, the early positivist view of reality is virtually indistinguishable from the view of modern qualitative research, as is evidenced by the statement below by one of the Vienna Circle’s early members in discussing the presuppositions of the Vienna Circle’s program:

One kind of observation is not truer or more faithful to reality than is the other, but the contexts in which they occur differ and must be described by different words. Every scientific statement is a statement about complexes of sensations, and beyond of behind these there are no realities to be looked for, because the word “reality” itself is merely a name for the sum total of the complexes of observable sensations. (Joergensen, 1951, p. 9)

Like positivists, many qualitative researchers believe that their observations are theory-free and that their theories will emerge from the data (e.g., Lincoln & Guba, 1985) through inductive processes. On this point, however, both positivists and qualitative researchers are incorrect. In countering this positivist assertion and view, the philosopher Hanson (as cited in Eldredge & Gould, 1972) contends that “much recent philosophy of science [i.e., postpositivism] has been dedicated to disclosing that a ‘given’ or ‘pure’ observation language is a myth eaten fabric of philosophical fiction . . . . In any observation statement the cloven hoof-print of theory can readily be detected” (p. 85). The idea that we observe, perceive, analyze, and make judgments in the natural world (and in professional and research situations) largely in relation to our existing or dominant theoretical constructs most of the time is extensively substantiated in the experimental findings of modern cognitive scientists and psychologists across diverse fields and is hardly a debatable point (see Ashcraft, 2002). Most notably in this regard is the well-documented and embarrassing variability and high degree of unreliability of psychiatric disease diagnoses between psychiatrists prior to the development of agreed upon and published diagnostic criteria (pre-1980s), which was linked to different theoretical schools of thought (Rosenhan, 1973).

It is exactly these types of intractable issues and problems that are ameliorated and neutralized when we think of research and evaluation programs as existing along a continuum, moving or shifting between the context of discovery (versus qualitative research) toward the context of justification (versus quantitative research). The four problems outlined in this section also suggest that trivializing a philosophical movement and adopting an *extreme* view or segment of such movements to support a particular intellectual endeavor (i.e., positivism and the context of justification or postpositivism and the context of discovery) is not simply disingenuous and

lazy scholarship, but a distortion of the actual nature and intent—and therefore a distortion of the actual virtues and limitations—of each program, or what we might call *justification error*. Parenthetically, we see a similar form of distortion and threat to validity when we sample *extreme* groups or data sets in educational testing environments due to regression effects and measurement error. In other words, measurement error and validity threats are not just psychometric issues, but ubiquitous phenomena that can also be applied to how researchers “fish from” and select aspects of theoretical and epistemological referents and apply them to their own research (see Carifio, 2005). Also parenthetically here, our conceptions of validity cannot be relative to methodologies let alone theories (see Carifio & Perla, 2008, for further explication of this point) and research findings that are truly unique (which *is* a statistical concept) are of little value to anyone unless, of course, one is trying to develop and confirm a predictive theory of unique events or a method of empirically identifying findings and events that one can safely ignore and exclude from consideration, as has been done in astronomy (see Galison, 1987, for details). Neither of these claims of qualitative theorists can withstand even minimal philosophical or logical scrutiny.

The points that are being emphasized here are that (1) logical positivism did indeed involve a nonexperimental and nonlogical context of discovery phase that led to the crystallization of certain ideas and commitments that were capable of being tested and falsified and (2) both the context of discovery and the context of justification are needed for a complete view and synthesis of any field of study looking to describe knowledge production and research in a meaningful way. As Fleck’s (1935) classic work suggests, we move from highly speculative claims and metaphysical assertions (e.g., syphilis is a punishment from God) toward claims that are made clear and then tested to determine their value (e.g., syphilis is caused by the

bacterium *Treponema pallidum*). Once a set of ideas have been tested and selectively retained in a particular knowledge base as the rudiments to theory, future observations and tests determine how and in what direction the knowledge base develops (Suppe, 1974). And as stated above, this process (movement between the two contexts) can and often does go “backwards.” The context of discovery is associated with fuzzy, generative, creative, intuitive, emotive, and informal thought, whereas the context of justification is associated with the formal testing and falsification procedures used in science (such as statistical hypothesis testing and regression analysis). Both the context of discovery and the context of justification are needed in nuanced and sophisticated research, and understanding and studying the relation between these contexts can inform the current debate in educational research methodology.

## General Research Considerations

Having outlined an epistemic framework and argument for a more comprehensive and complete view and model of research that transcends the qualitative and quantitative paradigms, this section of the paper outlines eight general key points that most (if not all) researchers and evaluators should consider, particularly educational researchers and researchers in the social sciences. Consistent with the model of Kleinig (1982), our list is intended to begin a process that identifies important research issues, concepts, and problems that are general enough to establish agreement on *basic and inclusive* research principles (i.e., heuristics). Each of these points is grounded in the philosophical and epistemological positions advanced earlier in this paper and serves as a summary of the main ideas and positions. These eight points are also interrelated to some degree and have many subelements that define them. Only a brief description of the eight points is provided

below to give a general sense of how these points are applicable, at least conceptually, to most research and evaluation endeavors. A fuller elaboration of these points and the addition of other points is the focus of ongoing study.

1. Research and evaluation involves the collection, analysis, interpretation, and organization of data (measurements, observations, evidence) of some type. Given that research is data-driven, then one must strive to have the best data possible given the context and situation at hand, and to use the best, most appropriate data analysis procedures available.
2. Observations in research and evaluation are almost always theory-driven and made meaningful by being compared to some expectation. Researchers in the social and educational sciences (like all researchers) usually undergo extensive training and schooling that implicitly or explicitly addresses theoretical issues and concepts. The better we recognize, understand, and develop our existing theoretical knowledge base, the better equipped we are to develop our (theory-based) expectations that can be measured against direct experience.
3. All research includes measurement and/or observation error, and the better we control for and understand measurement error, the more we understand the nature of our findings. The second a researcher or evaluator decides to measure (or observe) something, measurement error is introduced—regardless of what we are trying to measure (and especially in the social and educational sciences). A qualitative researcher looking to tell a story based on an interview, for instance, has the possibility of “errors in interpretation” (and many other potential sources of error, depending how controlled the observations are). As mentioned by Reichenbach (1938) above, measurement error in science is a well-recognized problem and much effort in science is aimed at reducing sources of error and invalidity or dealing with error that we know exists, but may be beyond our control. Regardless of the type of methods we use to collect, analyze, and index data, measurement error must be a major consideration in how we address the scope, virtues, limitations, and generalizations of our data.
4. All research and evaluation is a meaning-making system, even if implicitly. It is helpful to think of research as a meaning-making system since our ultimate goal is to share knowledge with others in our fields of study and to contribute meaningfully to and extend the knowledge base. Research is the vehicle by which new knowledge is added to (or deleted from) our collective understandings of certain problems and issues.
5. All research and evaluation is incomplete and tentative. The dominant view among all scholars, regardless of their field, is that knowledge, although reliable and extremely stable in some situations, is never complete, perfect, or indubitable. If this were not the case, it would be impossible for knowledge to grow and develop, because it could not change. Among the things we deal with in our human existence and especially in our evaluations and inquiries into human behavior are imperfection, error, shortsightedness, and confusion. However, knowing that the possibility of error exists should keep us open to new ways of thinking, especially in our research efforts as we shift between the context discovery and justification. In this sense, it is helpful to think of knowledge as “selectively retained tentatives.”
6. The research and evaluation method you use at one point in time does not define you as a researcher or a person. One of the major hurdles to a unified model of sound research and evaluation are the self-induced

boundaries and factions created by some who suggest that research is either qualitative or quantitative. Students need to be familiar and competent in the techniques associated with what is and has been described as qualitative and quantitative research, and qualitative research should not be the de facto standard for people unfamiliar with or uncomfortable with quantitative analysis. We all need to read and heed Frank Phillips' (2005) article on the need for professionals and practitioners who are fairly expert in conducting and evaluating instances of all types and cases of research and inquiry. Further, instead of referring to mixed methods, perhaps we should refer to shifting contexts and priorities—especially when exploring evaluation strategies over time.

7. Research and evaluation findings that can be tested, falsified, and replicated have the greatest value. This point is the case in all fields of study as researchers and society in general have more faith and confidence in theories, applications, and products that have successfully stood the tests of multiple falsification attempts over time (versus theories and products that have never been tested). As Popper notes, "Every genuine test of a theory is an attempt to falsify it, or to refute it . . . It is easy to obtain confirmations, or verifications, for nearly every theory—if we look for confirmations" (as cited in Schick, 2000, p. 11).
8. All research and evaluation in the social sciences and education implies a "theory of the responder," that is, a view of learning and memory that relates to the subjects in the study, which should be clearly identified and elaborated by the researcher up front. Research that involves interaction, communication, or assessment of people and their behaviors should involve an articulated view of the responder, as this is needed for the reader to understand the dynamics, design, results, and conclusions

of a study. For example, a researcher that espouses a radical behaviorist view of memory and learning will interpret data differently and draw different conclusions vis-à-vis a neobehaviorist who will draw different conclusions from a cognitivist or constructivist view or theory. An educational researcher who has not carefully considered a theory of the responder is not much different than a research chemist who has not developed a theory of the atom. Further, qualitative researchers often talk about emerging theory, yet without articulating their view of the responder, the emerging theory may be their existing view of the responder that was undefined or unrecognized from the beginning of the study.

## Conclusion

This paper addresses an important problem that may really be a pseudoproblem perpetuated by a current (artificial or even political) need to isolate researchers, evaluators, theorists, and professionals engaged in measurement and put them in conceptual-methodological boxes. In this paper, we argue that scholars of all kinds should focus on the nouns (research, evaluation, and theory) versus the adjectives (quantitative, qualitative or mixed-method). In other words, researchers, evaluators, and theorists of all dispositions should strive to be neither qualitative nor quantitative, as that is like, in our opinion, striving to be a single leg in a triangle. As stated at the beginning of this paper, our view is that the "patchwork quilt" (Hacking, 1965) that is both methodology and analysis is essentially all "one fabric" with two connected, but often unacknowledged, sides to the quilt; and all methodology and analysis is essentially "qual-quantification" and "quant-qualification" and usually both at the same time. This general and unified or "one-fabric-even-if-a-patchwork-quilt" view of these debates and issues was discussed and articulated in the context of

classic philosophical referents, and the need for a more general and inclusive research and evaluation epistemology was elucidated.

This paper ambitiously then developed an epistemic framework and philosophical justification for a more comprehensive and general view and model of nuanced and sophisticated research and evaluation and showed how this approach ameliorates intractable problems and logical inconsistencies in the isolationist paradigm and view that quantitative and qualitative methods are, or should be, viewed as mutually exclusive (and antithetical to each other). These epistemic and philosophic arguments were made concrete by identifying eight basic and inclusive points related to research methodology and evaluation. Our view that there exist nuanced, core, and inclusive principles of meaningful inquiry that transcend specific methodologies, which have almost become cult-like in their reifications, is critically important for students and future researchers and evaluators to understand if we are to substantially increase the quality of educational and social science research and develop professionals and practitioners who are fairly expert in conducting and evaluating instances of all major types and cases of research and inquiry.

### Authors' Note

1. The fundamental aspects of these debates are, of course, far from new and can be traced back in formal fashion to the work of Rene Descartes and his *Discourse on the Method*, published in 1637.

### Acknowledgements

Rocco J. Perla is currently a George W. Merck Fellow at the Institute for Healthcare Improvement, Cambridge, MA 02138.

### References

- Aneshensel, C. S. (2002). *Theory-based data analysis for the social sciences*. Thousand Oaks, CA: Pine Forge Press.
- Ashcraft, M. H. (2002). *Cognition* (3rd ed.). New Jersey: Prentice-Hall.
- Brass, C. T., Nunez-Neto, B., & Williams, E. D. (2006). *Congress and program evaluation: An overview of randomized control trials (RCTs) and related issues*. Retrieved October 24, 2008, from <http://digital.library.unt.edu/govdocs/crs/permalink/meta-crs-9145:1>
- Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research*, 1, 8-22.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 54, 81-105.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin.
- Carifio, J. (2005, July). *Towards a standard integrated information processing/cognitive model of learning*. Paper presented at the meeting of the International History, Philosophy, and Science Teaching Group, Leeds, England.
- Carifio, J., & Perla, R. J. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, 42, 1150-1152.
- Carnap, R. (1937). *The logical syntax of language*. London: Kegan Paul.
- Coryn, C. L. S. (2007). The 'holy trinity' of methodological rigor: A skeptical view. *Journal of MultiDisciplinary Evaluation*, 4(7), 26-31.
- Creswell, J. W. (1994). *Research design: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.
- Denzin, N., & Lincoln, Y. (Eds.). (2005). *The Sage handbook of qualitative research*. Thousand Oaks, CA: Sage.



- Eldredge, N. & Gould, S. J. (1972). Punctuated equilibria: An alternative to phyletic gradualism. In J. M. Schopf (Ed.), *Models in paleobiology* (pp. 82-115). San Francisco: Freeman, Cooper.
- Fleck, L. (1979). *Genesis and development of a scientific fact* (F. Bradley & T. Trenn, Trans.). Chicago: University of Chicago Press. (Original work published 1935)
- Frank, P. (1949). *Modern science and its history*. New York: George Braziller.
- Galison, P. (1987). *How experiments end*. Chicago: Chicago University Press.
- Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review*, 98, 254-267.
- Glesne, C. (1999). *Becoming qualitative researchers: An introduction*. New York: Addison Wesley Longman.
- Hacking, I. (1965). *The logic of statistical inference*. Cambridge: Cambridge University Press.
- Hanson, N. R. (1958). *Patterns of discovery*. Cambridge: Cambridge University Press.
- Helmer-Hirschberg, O., & Rescher, N. H. (1960). *On the epistemology of the inexact sciences* (Rand Corporation Document No. R-353). Retrieved November 24, 2008, from <http://www.rand.org/pubs/reports/R353/>
- Joergensen, J. (1951). The development of logical empiricism. In *International Encyclopedia of Unified Science*, (Vol. II, No. 9). Chicago: University of Chicago Press.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed method research: A research paradigm whose time has come. *Educational Researcher* 33, 14-26.
- Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research*. Orlando, FL: Harcourt.
- Kleining, G. (1982). *An outline for the methodology of qualitative social research*. Hamburg: University of Hamburg. Retrieved October 22, 2008, from <http://www1.uni-hamburg.de/abu//Archiv/QualitativeMethoden/Kleining/KleiningEng1982.htm>
- Kraft, V. (1953). *The Vienna Circle: The origins of neopositivism. A chapter in the history of recent philosophy*. New York: Philosophical Library.
- Kuhn, T. S. (1970). Logic of discovery or psychology of research? In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 1-23). Cambridge: Cambridge University Press.
- Kuhn, T. S. (1962/1996). *The structure of scientific revolutions* (3rd ed.). Chicago: University of Chicago Press.
- Lakatos, I. (1970). Falsification and the methodology of scientific research programs. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 309-330). Cambridge: Cambridge University Press.
- Lincoln, Y. S., & Guba, G. (1985). *Naturalistic inquiry*. Thousand Oak, CA: Sage.
- Matthews, M. R. (2004a). Thomas Kuhn's impact on science education: What lessons can be learned? *Science Education*, 88, 90-118.
- Matthews, M. R. (2004b). Reappraising positivism and education: The arguments of Phillip Frank and Herbert Feigl. *Science & Education*, 13, 7-39.
- McCall, W. A. (1923). *How to experiment in education*. New York: Macmillan.
- Pawson, R., & Tilley, N. (2007). *Realistic Evaluation*. Thousand Oaks, CA: Sage.
- Perla, R. J. (2006). *Use and augmentation of a formal model and theory to develop instructional materials to teach undergraduates about the nature of science, scientific knowledge and scientific change*. Unpublished doctoral dissertation, University of Massachusetts, Lowell.
- Phillips, F. (2005). The contested nature of empirical research (and why philosophy of education offers little help). *Journal of Philosophy of Education*, 39, 577-597.
- Pinker, S. (1997). *How the mind works*. New York: W. W. Norton.
- Quine, W. V. O. (1953/1961). *From a logical point of view*. New York: Harper & Row.
- Reichenbach, H. (1938). *Experience and Prediction*. Chicago: University of Chicago Press.

- Rosenhan, D. (1973). On being sane in insane places. *Science*, 179, 250-258.
- Ryle, G. (1949). *The concept of mind*. Chicago: University of Chicago Press.
- Schick, T. (Ed.) (2000). *Readings in the philosophy of science: From positivism to postmodernism*. Mountain View, CA: Mayfield.
- Shavelson, R. J. & Towne, L. (Eds.) (2002). *Scientific research in education*. Washington, DC: National Academy Press.
- Sloane, F. (2008). Through the looking glass: Experiments, quasi-experiments and the medical model. *Education Researcher*, 37, 41-46.
- Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models and applications*. San Francisco: Jossey-Bass.
- Suppe, F. (1974). *The structure of scientific theories*. Urbana: University of Illinois Press.
- Thagard, P. (1988). *Computational philosophy of science*. Cambridge, MA: MIT Press.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice, *Science*, 211, 453-458.
- Wittgenstein, L. (2003). *Tractatus logico-philosophicus*. (C. K. Ogden, Trans.). New York: Barnes & Noble Books. (Original work published 1922)