Journal of Theoretical and Philosophical Psychology
2016, Vol. 36, No. 4, 199–213

1068-8471/16/$12.00 http://dx.doi.org/10.1037/teo0000035

Reconsidering Philosophy of Science Pedagogy in Psychology: An Evaluation of Methods Texts

Joshua W. Clegg
John Jay College and City University of New York

Some have suggested that philosophy of science pedagogy in psychology is largely positivist and anachronistic. The present work assesses that pedagogy via a detailed analysis of all philosophy of science assertions made in 9 commonly used research methods texts. The resulting analysis describes a generally consistent philosophy of science narrative across all of the texts analyzed. That narrative is defined by the ontological assumptions of an independent and determinate reality; the related assumptions of empiricism, fallibilism, objectivism, and universalism; and the epistemological practices built on these assumptions—namely, operationalism, quantification, reliability and validity, and falsificationism. This analysis does, indeed, suggest a generally positivist philosophy of science pedagogy and so accords substantially with existing critiques, though some treatments, including those of realism and of multiple methods, showed more contemporary readings of philosophy of science. Following this analysis is a brief discussion of the possible relevance to psychology of reforms, undertaken in general science education, whose intent has been to make science teaching more current with respect to philosophy of science. These reforms suggest that concepts and scholarship from the history and philosophy of science could be used to frame research methods, such that these can be presented, not as abstract and independent truth procedures, but as contested and complex histories of socially and politically embedded practices.

Keywords: methods, philosophy of science, textbooks, pedagogy, psychology

There are some things evident to anyone familiar with philosophy of science that are nevertheless baldly contradicted in research methods courses across the psychological disciplines; and probably quite often by those of us who know better. I take a slow, deep breath every time I teach students to construct an operational definition. In the process, I try to tell them that the notion of an operational definition was a short-lived prescription from a short-lived philosophy of science (short-lived in philosophy anyway; Bickhard, 2001); that it is a confounding and nonsensical idea (Green, 2001) but psychologists stick to it anyway. But this experience is profoundly unsatisfying to all of us; my students want to know the "right" way to do research and I do not know how to explain to them that the science concepts and practices they need to succeed in graduate school and on the job market do not constitute a coherent philosophy of science; that if I taught them precisely what their discipline wants them to know about research, they would have acquired a naïve, antiquated, inaccurate, and confused picture of science theory.

The problem of an inadequate philosophy of science pedagogy is not unique to psychology. In an analysis of general science texts, Niaz (2010) found that "textbooks are 'systematically misleading' our future generations of citizens and scientists, by providing an ahistoric account of progress in science in which presuppositions, contradictions, conflicts and controversies have no role to play" (p. 896). Rodriguez and Niaz (2004) drew a similar conclusion from an analysis of physical science textbooks,

Editor's Note. Kate Slaney served as the action editor for this article.—BDS

This article was published Online First May 5, 2016.

Correspondence concerning this article should be addressed to Joshua W. Clegg, Psychology Department, John Jay College, 524 West 59th Street, New York, NY 10019. E-mail: jclegg@jjay.cuny.edu

arguing that these ignore the inconsistencies and heuristics in science work. This account is not far from Kuhn's (1996) assessment of science texts whose purpose, as Costa and Shimp (2011) described it, is "to persuasively argue the charms of a science and to suppress conflicting information" (p. 28).

This pedagogic deficiency has not gone unacknowledged and, over the last 20 years, reforms in general science education have attempted to address the gaps between science teaching and science theory and practice. These reforms have emphasized the social relevance of science (Santos, 2009; Zembylas, 2006), as well as the importance of discovery and explanation, as opposed to rote learning (Braaten & Windschitl, 2011). Under these reforms, "many modern science curricula place great emphasis on conveying science as a dynamic, human endeavor" (Pearce, 2013, p. 532) and this has pointed toward a highly contextualized and historicized science curriculum:

> Many of the recent proposals regarding the philosophy of contemporary science education suggest that a new paradigm of science education is required; namely, one that takes into account the postpositivist tendencies of the philosophy of science, which emphasize the contingent, linguistic, and historical nature of scientific knowledge and pay attention to science's underlying interests, power struggles and interconnections with the lifeworld. (Leiviskä, 2013, p. 516)

Some of these proposals are somewhat radical—for example, those that are explicitly critical (Santos, 2009) or postmodern (Zembylas, 2000)—but even in the mainstream of science education there is "a fair amount of consensus . . . for the need to write textbooks within a historical perspective" (Niaz, 2010, p. 892).

General science education, then, has begun to see the importance of integrating contemporary scholarship from history and philosophy of science. Unfortunately, we see very little of that spirit of reform in psychological science pedagogy. This is undoubtedly due in part to the fact that science education is an integrated discipline with national associations and best practices, while philosophy of science education in psychology (like our history education) is still considered a "gentleman's science"—that is, we seem to assume that it requires no special training and can be taught, primarily as part of a methods course, by any professional psychologist. Thus those who teach philosophy of science concepts to psychology students may have no formal connection to, nor a contemporary reading of, those who produce philosophy of science scholarship. It would not, then, be surprising if (as has been the case in other sciences) our philosophy of science pedagogy is overly simplistic.

But a positive critique of psychology's philosophy of science pedagogy is somewhat beforehand. Such a critique, and perhaps some alternatives, awaits an assessment of precisely what sort of philosophy of science is actually being taught to psychology students. No doubt we all have our own experiences with methods instruction (perhaps similar to the descriptions offered above), but addressing the question of pedagogy reform requires a somewhat more comprehensive analysis than personal experience can afford. This more comprehensive account is my purpose here; specifically, I explicate something of the philosophy of science being taught to psychology students by way of an in-depth thematic analysis of commonly used research methods texts.[1]

There is some precedent for taking this approach. Two recent articles have attempted to interrogate elements of psychology's philosophy of science pedagogy through an analysis of methods texts. In the first, Proctor and Capaldi (2001) assessed the extent to which commonly used texts reflected contemporary philosophy of science regarding hypothesis testing. Their analysis suggested that notions of hypothesis testing reflected little theoretical work since 1960 and that in "widely used texts . . . [this] coverage tends to be largely similar" (p. 173). In the second, Costa and Shimp (2011) analyzed the indices of commonly used methods texts for evidence of positivist terminology (e.g., "operational definition," "hypothesis testing," etc.), terms they found to be nearly ubiquitous. These conclusions were also corroborated by a recent analysis of psychology methods courses in

[1] The texts chosen were written by, and are used by, psychologists in various countries, but it should be acknowledged at the outset that this analysis centers the U.S. context. There are good reasons for this—U.S. psychology has had a disproportionate influence on psychology in general—and more circumstantial ones—this is the context within which I work and with which I am familiar. In any case, the present analysis should be considered U.S.-centric with all the limitations this entails.

South Africa, where "an examination of the content of the courses revealed that there is a heavy reliance on methods that are traditionally linked to the positivist paradigm" (Wagner & Maree, 2007, pp. 122–123).

These articles provide some clear evidence for a generally positivist philosophy of science in psychology's methods instruction (though not a coherently and explicitly defended sort of positivism) and this accords with general critiques of psychological theory (see review in the discussion below). These analyses, however, were rather narrowly focused and were essentially confirmatory tests of certain assumptions. My analysis here takes a similar methodological assumption—that is, that an analysis of textbooks can reveal something about science pedagogy in psychology—but my goal is to provide a more comprehensive account, based on essentially all of the philosophy of science assertions expressed in these texts. My hope is that this will produce a reasonably accurate account of the general philosophy of science espoused in these texts (and, with appropriate caveats, in the field more generally) and so provide some ground for specific and substantive critique.

To do more than give a reasonably complete outline of this analysis is not possible in a paper of this size, so any detailed critique of the philosophy of science represented in these texts will have to await further work. In my concluding remarks, however, I do offer some reflections on the general contention outlined above that philosophy of science concepts in psychology textbooks are largely positivist. I will also briefly consider whether philosophy of science education in psychology could benefit from those reforms, undertaken in science education, whose intent has been to make science teaching more current with respect to philosophy of science.

A Textbook Account of Psychology's Philosophy of Science

The following account is based on an analysis of commonly used methods texts, with the assumption that texts in broad adoption are good representations of typical methods instruction in psychology. The involved nature of the analysis required that I select a relatively small number of texts ($n = 9$) for in-depth analysis, but I do not believe this presents a serious challenge to the representativeness of the analysis presented. Like Proctor and Capaldi (2001), I surveyed a large number of methods texts used in psychology and found their philosophy of science coverage substantially similar; indeed, thematic "saturation" was reached well before completing the analysis of the nine texts chosen. Admittedly, "broad" or "common" use is not a completely straightforward notion, but I used what I consider reasonable, though certainly not exhaustive, selection criteria—namely, how often texts are purchased and how often they are assigned.

Specifically, I compiled a list of the best-selling research methods texts in psychology, one each from Amazon and Barnes and Noble (for those texts highest on these two lists, there was substantial overlap but very little in those texts lower on the list). Any books on these lists are likely read by thousands of students each year, and so they are good indicators of broad use. I also divided the United States into the four geographic regions identified by the U.S. census (Northeast, West, Midwest, and South), and in each region, ranked psychology departments based on the size of their faculty and on the average number of publications and citations for faculty members. I assumed that departments with large numbers of faculty actively engaged in research would be likely to teach comparatively large numbers of research methods courses (and students). I then compiled a list, where possible, of the required texts in undergraduate and graduate psychology methods courses for the two highest ranked public and the two highest ranked private universities in each region. The texts most often appearing on these lists were often the same best-selling texts from the Amazon and Barnes and Noble lists; outside of these commonly occurring texts, however, there was only occasional overlap across the various lists. Finally, I compared all of these lists, which between them included 100s of textbooks, and chose for analysis those nine texts that appeared more than once (see Table 1 below for a list of the texts analyzed).

The texts were analyzed using an in-depth thematic approach. First, each text was read in its entirety and any passages with philosophy of psychological science implications were grouped according to a set of continually evolving themes. For the purposes of this analysis,

Table 1
Research Methods Texts Analyzed

Text analyzed
Cozby, P. C. (2009). *Methods in behavioral research*. Boston, MA: McGraw-Hill Higher Education.
Gravetter, F. J., & Wallnau, L. B. (2009). *Statistics for the behavioral sciences*. Belmont, CA: Wadsworth.
Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research*. Fort Worth, TX: Harcourt College Publishers.
Martin, D. W. (2008). *Doing psychology experiments*. Belmont, CA: Thomson/Wadsworth.
Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.
Pelham, B. W., & Blanton, H. (2007). *Conducting research in psychology: Measuring the weight of smoke*. Belmont, CA: Thomson Wadsworth.
Rosnow, R. L., & Rosenthal, R. (2008). *Beginning behavioral research: A conceptual primer*. Upper Saddle River, NJ: Pearson/Prentice Hall.
Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2009). *Research methods in psychology*. Boston, MA: McGraw-Hill Higher Education.
Sommer, R., & Sommer, B. B. (2002). *A practical guide to behavioral research: Tools and techniques*. New York, NY: Oxford University Press.

any assertions that either included general statements about the nature or purposes of science or that served to justify or evaluate, rather than merely describe, particular methodological practices were considered expressions of some kind of philosophy of psychological science. Such a broad inclusion criterion was necessary because none of the texts reviewed here were either elaborating or advocating a specific philosophy of science. Many texts covered some philosophy of science concepts and history (generally at the front), but the methodological prescriptions occupying the majority of the texts[2] were not explicitly grounded in any particular philosophy of science. Rather, the philosophy of science assertions included in the analysis below were almost entirely offered in the context of general statements about science or research and were not specific to any particular method.[3] Though these assertions certainly do not constitute a specific philosophy of science, they do represent a kind of “folk” philosophy of science—that is, a generally accepted set of assumptions about what science is and how it works that was remarkably consistent across all of the texts analyzed here.

On the initial read of each text, I underlined any passages consistent with the above criteria and, in a separate document, noted down the page number and a summary of the passage. These summaries were then compared to identify common themes and to generate a logical structure for the general theory of science reflected in the texts. Finally, these themes were organized into a general description of the philosophy of science represented in all of the texts and each underlined passage was reread in light of that general description. As in most thematic approaches, this was not a probabilistic analysis; only those themes consistent with every text analyzed were included in the final description. There is, however, one important caveat to this general approach—namely, one of the texts included in the list was a purely statistical text and so does not include explicit endorsements of many of the basic philosophy of science assumptions included in the other texts. I considered excluding this text but decided against that approach for two important reasons: (a) given that this book is, in fact, being used as the main text in multiple methods courses, it is important to include it in the analysis; and (b) though this text did not include explicit endorsements of every feature described in the analysis below, it is certainly consistent with, and does not in any way contradict, any of those features (in addition to actually endorsing many of them).

In the analysis that follows, representative quotes from the texts are provided as support for each aspect of the description being presented. For reasons of space, flow, and clarity, passages

[2] A quite consistent set, in the standard order of nonexperimental (with some variable, but generally scant, coverage of qualitative methods), then experimental, and finally quasi-experimental designs.

[3] Two quotes refer specifically to experimental designs; this is noted in the quotes themselves.

from only a few texts will be quoted as support for any given description. Quotes were, however, available from any of the texts (with the exception, in some cases, of the aforementioned statistics text) and I tried to represent all of the texts. The entirety of the analysis presented below is intended as a faithful description of the philosophy of science presented in the texts being analyzed and not as an argument for any particular philosophy or method. In this analysis, I made no attempts to either critique or justify any elements of the science theory being presented; rather, my only intent was to describe that set of assertions, and its associated logical framework, that was consistent across all of the texts. For every principle or practice detailed below, much more could have been said in the way of a full elaboration of each, but my purpose here was to describe the general philosophy of science content, not to justify or fully articulate it.

There are many assertions relative to the philosophy of psychological science that find broad endorsement across all of these texts, but nearly all of these assertions are grounded in two general ontological assumptions—namely, that reality or "the world" is (a) independent and (b) determinate.

Reality Is Independent

The clearest statement of the assertion that reality is independent of human understanding is found in Kerlinger and Lee (2000): "There are real things, whose characters are entirely independent of our opinions about them" (p. 7), or as Maxwell and Delaney (2004) put it, "realism holds that a definite world exists . . . and 'the way the world is' is largely independent of the observer" (p. 19). Rosnow and Rosenthal (2008) make the same point in a slightly different way: "This is the role of the scientific method, to provide a framework for drawing on independent realities to evaluate claims rather than to depend on tradition, authority, or armchair reasoning" (p. 6). This ontological thesis provides the grounding for two basic epistemological assumptions—because the world is independent of our understanding of it (and not constructed by, or identical with, that understanding): (a) we must turn to our senses for direct knowledge of that world (i.e., empiricism)[4] and (b) errors of interpretation are always possible and so there will always be uncertainty about the relationship between interpretation and reality (i.e., fallibilism).

Empiricism. The empiricist assumption is probably the least controversial across all of the texts. As Cozby (2009) put it, "knowledge is based on observations" (p. 5). This assumption derives from the notion that reality and our understanding of it are not identical: "subjective belief, in other words, must be checked against objective reality" (Kerlinger & Lee, 2000, p. 14). Good scientists, then, rely on "empirical evidence rather than their subjective judgment" (Shaughnessy, Zechmeister, & Zechmeister, 2009, p. 15) and, among the various kinds of knowledge, "weight should be given to empirically based inductive arguments" (Maxwell & Delaney, 2004, p. 22).

Fallibilism. Under the assumption of fallibilism, all inferences, even those drawn from empirical reasoning, are limited, fundamentally uncertain, and "ultimately probabilistic" (Shaughnessy, Zechmeister, & Zechmeister, 2009, p. 17)—"we can never conclude anything with absolute confidence" (Pelham & Blanton, 2007, p. 61). The primary consequences of fallibilism are (a) that scientists must accept that there is "a certain amount of ambiguity in the enterprise of scientific inquiry" (Cozby, 2009, p. 9), that "all causal relationships are contingent or dependent on the context" (Maxwell & Delaney, 2004, p. 8) and so "scientific knowledge is relative and provisional" (Rosnow & Rosenthal, 2008, p. 3); and (b) that scientific knowledge is based on the "weight of evidence" (Kerlinger & Lee, 2000, p. 29)—that is, on an accumulation of empirical tests "accurate at only a probabilistic level" (Martin, 2008, p. 57), rather than on a definitive proof, or logically necessary demonstration of some theoretical principle.

Objectivism. The assumption of an independent reality, then, constructs scientific activity as an approximation, via the fallible activities of empirical reasoning, of a never fully known world. In the texts analyzed here, this vision of scientific activity is expressed through

[4] In addition to the assumption of an independent reality, empiricism also requires the assumption that the physical senses have a privileged access to this reality, as discussed in subsequent sections.

a sort of technical praxis that can best be labeled objectivity or objectivism (the term by which it is most often called). I call objectivism a technical praxis because it is not just an assumption; it is a large and complex tradition of practices aimed at limiting the ambiguity or uncertainty inherent in the inferential process.

Objectivist practice is based on the assumption that the fallibility of human reasoning injects uncertainty into science and thus the degree to which we can extract the process of inference from biased individual reasoning and relocate it in an independent and intersubjectively validated system of inference, the more valid our inferences will be. In the texts analyzed here, fallibility is generally conceptualized in terms of bias, a very broad concept that includes all of the ways in which the errors of human reasoning can produce unintentional effects in the design, execution, or interpretation of empirical research. Many forms of bias are discussed across the texts, with the most elaborate discussions focusing on (a) instances where the design or execution of research unintentionally alters the behavior of research subjects and (b) instances where the thoughts, feelings, or expectations of the researcher unintentionally alter his or her interpretations of subject behavior, or where "patterns observed in a set of data are influenced by the ideas the investigator brings to the research" (Maxwell & Delaney, 2004, p. 4). Grouped under the first category are phenomena like reactivity, demand, and observer effects—that is, "bias introduced by . . . human interaction" (Sommer & Sommer, 2002, p. 129)—the underlying assumption in each being that "the act of studying people can dramatically change the way that people behave" (Pelham & Blanton, 2007, p. 126). Grouped in the second category are phenomena like experimenter bias, observer bias, or interviewer bias, all of which involve researchers who "see what they are looking for" (Cozby, 2009, p. 134). These and a myriad of other possible individual biases constitute the inherent uncertainty of the inferential process and the praxis of objectivism is designed to circumvent, or at least limit, their influence by locating that inference "as much as possible in [the] reality lying outside the scientist's personal beliefs, perceptions, biases, values, attitudes, and emotions" (Kerlinger & Lee, 2000, p. 7).

That "reality" lying outside of scientist belief is primarily defined in terms of intersubjective practices—that is, in terms of the "publicly observable" (Martin, 2008, p. 143). "[The] empirical . . . is demonstrable; that is, it can be shown to other people" (Sommer & Sommer, 2002, p. 3). Intersubjective validation is thus the essential practice of objectivity: "agreement among a number of observers provides the basis for psychological measurement" (Shaughnessy, Zechmeister, & Zechmeister, 2009, p. 38). Kerlinger and Lee (2000) stated this even more explicitly: "objectivity is agreement among 'expert' judges" (p. 7). Phrases like "direct observation," then, seem to refer, not to a sort of direct apperception of the "really real," but to observational contexts where "others can replicate the methods used and get the same results" (Cozby, 2009, p. 5). The basic objectivist assumption, then, is that rigorous and skeptical communal validation will serve as a check against the vagaries of individual and biased reasoning. In short, "the cornerstone of the scientific perspective is openness to criticism and revision" (Pelham & Blanton, 2007, p. 21) where scientists "can report their research findings and these can be evaluated by others" (Cozby, 2009, p. 5) who "establish whether the research follows the rules of experimentation" (Martin, 2008, p. 96).

But, of course, any such validation must be conducted by individuals, the very sources of error that objectivist science is intended to control. The objectivist solution to this problem is found in method or system. Objectivist science "provides an objective set of rules for gathering, evaluating, and reporting information" (Cozby, 2009, p. 6) and these rules, themselves, become the ultimate arbiters of scientific fact—"scientific objectivity inheres in methodological procedures characterized by agreement among expert judges" (Kerlinger and Lee, p. 708). The simplest way to express this argument is in the assertion that science is "self-correcting" (Shaughnessy, Zechmeister, & Zechmeister, 2009, p. 186), such that "hypotheses exist, can be tested, and can be shown to be probably correct or incorrect apart from a person's values and opinions (biases)" (Kerlinger & Lee, 2000, p. 27).

Through the application of systematic method, then, science allows us to circumvent human bias and build independent knowledge, such that "each generation of researchers builds

on the important findings of previous researchers in a chain of discovery and understanding" (Rosnow & Rosenthal, 2008, p. xiv). In this sort of vision, science is seen as progressive and cumulative in a way that no other form of knowledge production is. As Pelham and Blanton (2007) claim, "the one hundred or so years in which people have conducted systematic research on human behavior have taught us more than we learned in the hundreds of centuries that preceded the last one hundred years" (p. 4).

Reality Is Determinate

The second basic ontological assumption running through these texts is that reality, unlike our knowledge about it, is determinate—that is, because "all events have meaningful, systematic causes" (Pelham & Blanton, 2007, p. 8), "the events of nature display a certain lawfulness" (Maxwell & Delaney, 2004, p. 6). The obvious extension of this assumption is that apparent disorder, or randomness, is an artifact of observation and not a property of the universe; "randomness is a term for ignorance" (Kerlinger & Lee, 2000, p. 182). "Randomness means that there is no known law, capable of being expressed in language that correctly describes or explains events and their outcomes" (Kerlinger & Lee, 2000, p. 167), but were we to fully understand all phenomena, such randomness would disappear—"to an omniscient being, there is no randomness" (Kerlinger & Lee, 2000, p. 182).[5]

Because randomness is assumed to be a reflection of researcher ignorance, "scientists refer to this as random variability or error variance" (Cozby, 2009, p. 71). Randomness is seen as "error" because it is the result of poor conceptualization, design, or measurement—that is, "chance fluctuations, or haphazard errors" (Rosnow & Rosenthal, 2008, p. 126)—and is not related to the qualities of the phenomena measured. Such error is considered "noise" because it obscures the ordered patterns that are the true goal of scientific investigation: "as the error variance increases, it becomes more difficult to see any systematic differences or patterns that might exist in the data" (Gravetter & Wallnau, 2009, p. 128). The function of good research design, then, is to minimize the "noise," or errors of measurement, in order to detect ordered and determinate patterns; the "consistencies in the data [that] will be discernible through the cloud of random variation" (Maxwell & Delaney, 2004, p. 8). In the texts analyzed here, this determinacy is encoded in various ways, the four most basic ones being (a) universalism, (b) reliability, (c) representativeness, and (d) convergence.

Universalism. First, because reality is seen as determinate, "research is aimed at reducing random variability by identifying systematic relationships between variables" (Cozby, 2009, p. 71). For scientists, nonpatterned or contingent relationships do not describe those "laws that allow them to make precise predictions of human behavior" (Pelham & Blanton, 2007, p. 28). Instead, these texts argue, scientists assume that "nature is uniform—that is, processes and patterns observed on only a limited scale hold universally" (Maxwell & Delaney, 2004, p. 6). One of the principle goals of scientific endeavor is thus to prune those theories that produce highly contingent or dis-ordered predictions and replace them with "general laws" (Sommer & Sommer, 2002, p. 9) that approximate perfect prediction: "Psychology (like science in general) seeks to establish broad generalizations, 'universal laws' that will apply to a wide population of organisms" (Shaughnessy, Zechmeister, & Zechmeister, 2009, p. 287). Strong scientific theories produce inferences with this generalizable, or law-like, quality, a quality generally referred to as external validity. Externally valid inferences are those that "will stand up over time" and that are "generalizable across different kinds of participants and different investigators" (Rosnow & Rosenthal, 2008, p. 137). The ideal scientific theory, then, is one that perfectly predicts in all contexts and so is perfectly isomorphic to an ordered and determinate universe.

Reliability (replication). A second way that the assumption of a determinate reality is reflected in these texts is through discourses around reliability. Reliability is discussed in two basic ways—first, in a fairly technical sense restricted to the context of measurement and, second, in a more broad sense referring to the

[5] Under this account of reality, determinacy is not a variable state (i.e., real objects having more or less definite properties) but an inherent quality (i.e., the nature of reality is such that objects have fixed, ordered, and knowable properties, even if these are not, at present, fully known).

general replicability of research findings. In the more technical sense, "a measurement procedure is considered reliable to the extent that it produces stable, consistent measurements" (Gravetter & Wallnau, 2009, p. 524); if "we get exactly the same result when we repeat the measurement a number of times under comparable conditions" (Martin, 2008, p. 137). Under this account, "stability" and "consistency" are epistemologically privileged such that unreliability indicates errors in measurement rather than indeterminate properties in the world.

In the more general sense, reliability is discussed in terms of the stability, or replicability, of research findings—for example, "the best way to determine whether findings obtained in an experiment are reliable (consistent) is to replicate the experiment and see if the same outcome is obtained" (Shaughnessy, Zechmeister, & Zechmeister, 2009, p. 202). The assumption here is that for an inference to be generalizable, or to approximate universal law, it must be repeatable across time and context. Consequently, multiple mutually consistent empirical tests of the same hypothesized relationship increase our confidence in the generalizable nature of that hypothesis.

Convergence. Reliability is closely related to the notion of convergence, which extends the logic of replication to other methods. The idea of convergence is that "when multiple studies using multiple methods all lead to the same conclusion, our confidence in the findings and our understanding of the phenomenon are greatly increased" (Cozby, 2009, p. 85). The assumption that stable findings across varied methods reflect an understanding of a phenomenon follows, again, from the notion that instability is error or noise of some kind. A natural corollary of the idea of convergence is that "it is essential not to foreclose on the use of tools and techniques that enable us to study phenomena from more than one relevant vantage point and to triangulate on questions of interest" (Rosnow & Rosenthal, 2008, p. xiv). "Triangulation allows the researcher to pinpoint aspects of a phenomenon more accurately by approaching it from different vantage points using different methods" (Sommer & Sommer, 2002, p. 59) and so a kind of multimethod, or "toolbox," approach to research is advocated in all of these texts (though more strongly in some than others).

Representativeness. A final expression of determinacy is the idea of representativeness. The basic assumption in discussions of representativeness is that the best theories reflect the ordered and definite nature of reality and so apply across all, or most, contexts. Just as reliability assumes repeatability across time and convergence across method, representativeness assumes repeatability across all comparable contexts: "by choosing times, settings, and conditions for their observations that are representative of a population of behaviors, researchers can generalize their findings to that population" (Shaughnessy, Zechmeister, & Zechmeister, 2009, p. 94). The fundamental assumption underlying this reasoning is that under comparable conditions, comparable effects can be expected to occur, an assumption that can only hold in a definitely ordered universe.

Under the broad assumption of determinacy, then, the goal of science is to derive law-like theories, and this lawfulness is demonstrated through a concatenation of mutually consistent empirical tests across multiple individuals, contexts, times, and methods. Because universal, or law-like, inferences must be consistent across all relevant contexts, it is never one test, one method, or one individual that can serve as evidence for generalizability, but only an aggregation of all relevant tests by all acceptable methods across as many individuals and contexts as possible (a notion we will return to shortly). Lawfulness, then, may be operable, or even determinable, under a single empirical test, but it can only be demonstrated in the aggregate.

Summary. To this point, my analysis has suggested that the philosophy of psychological science espoused in the texts analyzed here is based on two ontological assumptions: first, that reality is independent of human understanding, and so only knowable through the inherently fallible processes of empirical inference; and, second, that reality is determinate, and so only knowable as the ordered and consistent patterns we would expect from generalizable, or law-like, principles. From these assumptions are derived broad epistemological mandates—that is, general ideals for a scientific practice that will permit the justification of empirical inferences. The independence assumption is taken to justify the necessity of limiting, as much as possible, the degree to which individual interpretation governs scien-

tific inference by relocating that inference to independent systems of intersubjective validation. The determinacy assumption is taken as evidence for the importance of basing empirical inferences on the ordered and predictable patterns we find in mutually consistent observations aggregated across person, time, and context while distinguishing these from the inconsistencies, or noise, of error.

The texts analyzed here discuss a number of epistemological practices—that is, practices whose primary function is to justify empirical inferences—that are intended to fulfill these basic mandates. The four most central of these practices are: (a) formulating theories in terms of operational hypotheses that can be intersubjectively validated; (b) representing phenomena quantitatively; (c) validating the whole inferential process through a set of agreed-upon discourses and practices, known generally as reliability and validity, that codify the assumptions of objectivity and determinacy; and (d) evaluating the body of scientific inferences in terms of a probabilistic accumulation of evidence derived from the active falsification of hypotheses.

Epistemological Practices

Operationalism. Across the texts analyzed here, an *operational definition* is understood to be "a definition of the variable in terms of the operations or techniques the researcher uses to measure or manipulate it" (Cozby, 2009, p. 67). The basic assumption underlying this sort of operationalism is that "operations," or techniques for physical measurement or manipulation, exist "outside" the fallible interpretations of individuals and so operational definitions "help to clarify communication among scientists about their constructs" (Shaughnessy, Zechmeister, & Zechmeister, 2009, p. 36). "The primary reason for giving an operational definition is to specify clearly and precisely what is being measured so that the study can be repeated by someone else using exactly the same procedures and measurements" (Sommer & Sommer, 2002, p. 86) – that is, they create a ground for the intersubjective validation required by the assumption of an independent reality. On this logic, operational thinking "is the indispensable key to achieving objectivity (without which there is no science), because its demand that observations must be public and replicable helps to put research activities outside of and apart from researchers and their predilections" (Kerlinger & Lee, 2000, p. 59).

Representing phenomena quantitatively. A second major epistemological practice advocated in these methods texts is an insistence on the importance of drawing inferences from quantitative representations (i.e., counts or measures). There are at least three basic reasons for this practice: First, it is assumed that in scientific inference, "statistics, using probability theory and mathematics, makes the process more systematic and objective" (Kerlinger & Lee, 2000, p. 260). "Statistical analysis can objectify to some extent the process of looking for patterns in data" (Maxwell & Delaney, 2004, p. 5) and so "we can more rigorously test using quantitative designs" (Martin, 2008, p. 20); "instead of impressions about trends and biases, the investigator comes up with precise figures" (Sommer & Sommer, 2002, p. 178). Why quantities are considered more objective, rigorous, or precise is not made explicit.

The second major justification for drawing inferences from quantities is the assumption that statistical aggregation of multiple cases is necessary to demonstrate lawfulness. Psychological scientists, according to these texts, "should remain skeptical of conclusions drawn solely from a single case history" (Martin, 2008, p. 19). An individual observation, or even "a small number of observations is more likely than a large number of observations to reflect the operation of chance error" (Pelham & Blanton, 2007, p. 72) or of simple idiosyncrasy. A large number of observations, however, can, taken together (via statistical aggregation), constitute evidence for lawfulness: "other things being equal, large samples provide more reliable and representative data than small samples" (Sommer & Sommer, 2002, p. 240).

The final major justification for quantitative representations of empirical phenomena, and probably the most important, is the proposition that, under the logic of probability theory, random aggregation produces determinate patterns in sets of numbers. This proposition takes both a highly technical form and a more generalized one. The logic of the technical version is a bit too complex to elaborate here, but the basic idea (embodied in the Central Limit Theorem) is that, no matter the qualities of any given population of scores, the repeated random (or unbi-

ased) aggregation of samples from that population will, given enough repetitions, always be distributed in a determinate and definable way (i.e., in a Gaussian curve). This basic logic means that we have a mathematically definite model for random, or chance, aggregation and so through numerical aggregation (i.e., counting) psychological scientists can "set up chance expectations as their hypothesis and try to fit empirical data to the chance model" (Kerlinger & Lee, 2000, p. 258). The more general version of this argument is that all random aggregations will produce definite patterns and so, even if we cannot reliably predict individual events, with the help of random grouping "we can predict the outcomes of large numbers of events" (Kerlinger & Lee, 2000, p. 167). The general point here is that, in addition to the presumed objectivity of numbers and the necessity, imposed by the assumptions of universalism and fallibilism, of aggregating evidence in a probabilistic way, inference from quantitative representations is encouraged because that aggregation can be tested against the determinate patterns of random aggregation.

The discourses of reliability and validity. In the texts analyzed here, this whole process of testing theories and drawing inferences is overseen by a set of discourses and associated practices that embody the assumptions of objectivity and determinacy. These are constructed as independent rules for drawing proper inferences and so embody the systematic and externalizing approach to intersubjective validation. Many of these basic rules have already been discussed under their general names—reliability and validity—and so will only be considered here in their general character. The discourses of reliability essentially codify the assumption of determinacy across various contexts. Some forms of reliability focus on stability in test construction and others on the replicability of experimental findings, but all operate under the assumption that, in a determinate universe, law, or generalizability, will be reflected in repeated and consistent empirical results.

Validity (with some exceptions) is more generally focused on identifying and mitigating the errors of inference inherent in applying fallible individual reason to an independent reality. Validity, in its most general sense, is characterized as something very close to "truth": "validity means essentially truth or correctness, a correspondence between a proposition describing how things work in the world and how they really work" (Maxwell & Delaney, 2004, p. 23). When one has made a valid proposition, one has made the right, or true proposition—"the validity of a psychological statement refers to the relative accuracy or correctness of the statement" (Pelham & Blanton, 2007, p. 61). But "of course, because we can never conclude anything with absolute confidence, issues of validity are always relative rather than absolute" (Pelham & Blanton, 2007, p. 61). A valid inference, then, is an approximation, rather than a demonstration, of truth or accuracy.

Validity is generally approached from two quite different perspectives, one related to the design of research and the other to the assessment of research measures. The tradition of validity related to research design focuses on unknown or uncontrolled contextual influences that might lead the researcher to misinterpret the "true" relationships between phenomena. When attempting to ensure internal validity, the goal is to define and attempt to mitigate any such influences that may be present at the time and place where the research is conducted, so that "by having more control over the situation, the researcher can be more confident about what is being measured" (Sommer & Sommer, 2002, p. 4). The tradition of validity related to research measurement (e.g., content, criterion, and construct validity) focuses on gathering evidence to support the authenticity of some measure or representation of a research phenomenon—that is, to show that it "measures what it claims to measure" (Shaughnessy, Zechmeister, & Zechmeister, 2009, p. 38). Through this process, which is an intentional mirror of the general hypothesis testing approach to science, it is assumed that more accurate, or valid, forms of measurement are constructed.

Falsification. At the disciplinary level, the whole body of scientific inferences is also considered from the perspectives of determinacy and fallibility. The basic story for how knowledge is produced in the psychological sciences follows the same essential script across all the texts reviewed here: "scientists search for observations that will verify their ideas about the world" (Cozby, 2009, p. 5), but such observations are subject to the fallibility of empirical reasoning and, in any case, "we cannot anticipate all rival explanations" (Rosnow &

Rosenthal, 2008, p. 161), so "empirical tests can never prove that a theory is correct" (Pelham & Blanton, 2007, p. 33). Theories must, however, be tested—"if a hypothesis cannot be tested, it is not useful to science" (Shaughnessy, Zechmeister, & Zechmeister, 2009, p. 39). The solution to this problem is to rely on the assumption that though a theory cannot be proven empirically, "it is possible to prove that a theory is false" (Pelham & Blanton, 2007, p. 33) and so "our job as scientists is not to prove theories, but to disprove them" (Martin, 2008, p. 57).

Of course, even falsification is fallible—"not only is there no possibility of proving any scientific theory with logical certainty . . . there is no possibility of falsifying one with logical certainty" (Maxwell & Delaney, 2004, p. 15)—and so it is not one particular instance of falsification that builds scientific knowledge, but the whole history of falsification attempts across a discipline. The essential point of this story is to draw out the basic Popperian vision of science as "a free market of ideas in which the best ideas are supported by research" (Cozby, 2009, p. 6), such that "the fittest scientific formulations withstand falsification in what resembles an evolutionary process of competition and survival" (Rosnow & Rosenthal, 2008, p. 3). In this vision, science progresses as theories with poor predictive success are discarded in favor of those with good predictive success, and so, as Pelham and Blanton put it, "scientists do not travel forward toward truth so much as they walk backwards away from falsehood" (Pelham & Blanton, 2007, p. 33).

Summary and General Reflections

Though there was no commonly acknowledged theoretical framework invoked across the texts analyzed here, there was certainly a consistent philosophy of science narrative. That narrative was defined by the ontological assumptions of an independent and determinate reality, the related assumptions of empiricism, fallibilism, objectivism, and universalism and the epistemological practices built on these assumptions—namely, the discourses and practices of operationalism, quantification, reliability and validity, and falsificationism. A great deal of specific critique could follow on this account (and, indeed, I hope it does), but space forbids more than a brief consideration of the questions raised at the beginning of the paper. Specifically, I discuss below how the present analysis does, indeed, provide some support (with important caveats) for previous claims suggesting the broadly positivist and somewhat imbalanced philosophy of science coverage provided in methods texts; though, perhaps more importantly, it is the absence of any significant philosophy of science coverage that is most revealing. In addition, given that science education has already begun to address the problem of inadequate philosophy of science coverage, I briefly reflect on whether their efforts might provide some inspiration for reframing methods instruction in psychology.

Positivist?

The most common and generally agreed upon critique of psychology's philosophy of science pedagogy is that we are teaching a "mainstream" view of science that is broadly positivist and that our coverage of more contemporary philosophy of science is rather limited. This "mainstream" has been characterized as Western (Laungani, 2002), modeled on the natural sciences (Wertz, 1999), quantitative (Michell, 2003), experimental (Lynd-Stevenson, 2007; Tissaw & Osbeck, 2007), and, most commonly, positivistic (Faulconer & Williams, 1985; Burman, 1997; Martin, 2008; Kutney, 2006; Breen & Darlaston-Jones, 2010; or sometimes post-positivistic, Giddings, 2006, or neopositivistic, Luyten, Blatt, & Corveleyn, 2006). The "positivist" label is the most general descriptor, embodying essentially all of the putative characteristics of the mainstream. This label does not, of course, signify historical continuity with a particular positivist movement (of either Comtean or logical positivist variety) but is, rather, a kind of general "term of abuse" (Michell, 2003, p. 16). In critical accounts, the "positivist" worldview is associated with a methodological consensus that coalesced in the American psychology of the 1930s and 1940s. Elaborated by psychologists like S. S. Stevens and E. G. Boring (Michell, 2003; Martin, 2008; Danziger, 1997), the core assumption of this consensus was that "knowledge is objective and value free (or neutral), and is obtained through the application of the scientific method" (Breen & Darlaston-Jones, 2010, p. 68). A fair number of additional assumptions are ascribed to this

mainstream consensus, but the most central are universalism, naturalism, objectivism, realism, empiricism, experimentalism, and a bias toward quantification.

Some of these assumptions clearly framed the philosophy of science narratives offered in the texts analyzed here. Universalism, a perspective that "sees human action as a function of general laws relatively independent of time and space" (Faulconer & Williams, 1985, p. 1179), was certainly a basic assumption across all of the texts. The same can be said of a naturalism that "discard[s] all the metaphysical abstractions" (Laungani, 2002, p. 389; e.g., "metaphysical systems . . . were eventually abandoned by scientists" [Pelham & Blanton, 2007, p. 6]). In these texts "objectivism," or the attempt "to remove to as large a degree as possible the human condition, with all of its fallibilities, from the scientific method" (Costa & Shimp, 2011, p. 26) was also cast as an indispensable foundation for science.

Other assumptions generally imputed to the positivist mainstream are the realist assumption that there exists "an objective reality that is absolutely true, and [that] with the proper methods, reality may be known" (Kutney, 2006, p. 23) and the empiricist position "that privileges research as an 'empirical' rather than interpretive matter" (Burman, 1997, pp. 796–797). The empiricist assumption received unqualified endorsement in the texts analyzed here but discussions of realism were more nuanced than some critiques might suggest. It would certainly not be fair to characterize the realism espoused in these texts as uncritical or naïve—that is, these texts espoused the existence of an independent reality, but none of them claimed that scientists have an unproblematic or transparent access to that reality. All of the authors in these texts asserted the fundamentally fallible quality of scientific inference. Indeed, the fallibility of inference is precisely the reason for objectivist practice.

A final set of critiques attributed to the "positivist" mainstream are the belief that "the privileged method in psychology, the gold standard, is the experiment, in which alone a causally explanatory hypothesis . . . can be tested in a way that is, at least in principle, immune to extraneous influences" (Wertz, 1999, p. 133), and the quantitative bias that holds that "studying something scientifically means measuring it" (Michell, 2003, p. 6). Although these biases are clearly evident in the texts analyzed here, there is some evidence that they are not as entrenched or as extreme as some have claimed. There is in these texts a universal endorsement of multiple methods, including nonexperimental and nonquantitative approaches, though these approaches are generally still considered inferior to quantitative experimentation in some way. It is nevertheless revealing that in many of the texts, nonquantitative approaches are discussed in some detail and in their own terms. There is a kind of epistemological code switching where the "traditional" argument[6] is set next to an alternative[7] and the merits and limitations of both are discussed. So, though the quantitative and experimental biases remain, these are certainly not monolithic.

Generally, then, as Costa and Shimp (2011) and Proctor and Capaldi (2001) have argued, there are some clear elements of "positivist" philosophy of science in commonly used methods texts, though there is some question about whether "positivist" is obviously the best term. The term is referenced in very few of the texts and is never explicitly espoused in any of them as a guiding philosophy. Certainly in the broad colloquial way positivism is used as a term of abuse, the texts analyzed here reflect many of its tenets (e.g., objectivism, empiricism, operationalism, etc.), but none espouse an orthodox positivism, of either the Comtean or logical positivist kind.

Indeed, it cannot be said that any of the texts endorse any particular, integrated philosophy of science (with the possible exception of Maxwell and Delaney; though this is by no means clear). They all represent, instead, a bricolage of assumptions from a variety of sources, many of them not explicitly justified in terms of any particular theorist or tradition. The larger dilemma, then, is not whether psychology's "mainstream" philosophy of science pedagogy is positivist, but whether philosophy of science

[6] For example, "quantitative researchers argue that unless data can be converted into numbers, they can never be organized into the building blocks necessary for the construction of a scientific body of knowledge" (Martin, 2008, p. 19).

[7] For example, "qualitative researchers . . . say that quantitative researchers will never be able to understand realistic human behavior in a holistic way" (Martin, 2008, p. 19).

is, in fact, taken seriously in psychology's science education. Several of the texts, as Proctor and Capaldi (2001) noted in their analysis, referenced no ideas or theorists in philosophy of science more recent than Popper. Other texts did discuss more contemporary theorists, including Kuhn, Lakatos, Campbell, Meehl, Feyerabend, and Bhaskar,[8] but even in those texts with a more current treatment of philosophy of science, this was very brief (a matter of a few pages) and peripheral to the central concerns of research methods. There is no question that these texts provide far from adequate coverage of philosophy of science and, if this coverage is reflective of what psychology students generally receive (obviously a larger question), then we are guilty of the same simplistic philosophy of science pedagogy that has inspired many of the reforms in general science education.

Can Reforms in Science Education Inspire Reforms in Psychology's Methods Education?

There are undoubtedly many ways that we might address this inadequate philosophy of science coverage but, given that science education has been wrestling with this same problem for decades, it is worthwhile to briefly consider whether their efforts could offer some inspiration for psychology. As outlined in the introduction, reforms in general science education have attempted to "[take] into account the post-positivist tendencies of the philosophy of science which emphasize the contingent, linguistic and historical nature of scientific knowledge" (Leiviskä, 2013, p. 516). These reforms are thus not merely aimed at covering more recent philosophers of science; they are also designed to reframe science pedagogy in a way that reflects elements of post-positivist theory. What this means in practice is that "science education should encourage learners to question the role and functions of natural science in the context of the whole of human culture and society" (Leiviskä, 2013, p. 526). This contextualizing of science education involves historical elements—for example, using examples from science history (Pearce, 2013, p. 200) to show the "historical and linguistic contingency of scientific knowledge" (Leiviskä, p. 517)—as well as sociopolitical ones that promote critical thinking about the functions of science and that "help students engage in social issues" (Santos, 2009, p. 362). The overall purpose for this contextual approach is to cultivate the understanding that "science is not a dogmatic, eternal enterprise" but "an inherently human endeavor" (Pearce, p. 541).

Under this model, philosophy of science becomes less a content area, and more a perspective for making sense of science as practiced in a social context. Philosophy of science theorists, works, and concepts become pedagogical tools for framing the whole scientific enterprise. There are, of course, many other ways that psychology's philosophy of science coverage could be updated and, particularly for ideological opponents of the post-positivist turn, a focus solely on updated content, rather than on these more radical pedagogical reforms, would likely be more palatable. But for those interested in the kinds of innovations being explored in science education, I suggest a few possibilities for psychology.

One approach for integrating these more contemporary philosophy of science notions would be to restructure elements of a research methods course such that particular methods are taught as historical traditions, rather than as abstract and timeless scientific procedures. This would involve readings (from both history and philosophy of science) and discussions concerning how a method has been deployed and conceptualized across time, including attendant politics, controversies, ambiguities, and so forth. A similar, perhaps even complimentary, approach would be to juxtapose traditional coverage of a particular method or concept with prominent critiques and then create opportunities, in the classroom and in assignments, for dialogue and debate. One intriguing proposal by Santos (2009) is to identify some current social, political, or even practical, issues of immediate relevance to students and then engage with these using scientific practices and concepts, as well as important theoretical critiques.

In each of these approaches, the central concern would be to cultivate in students a critical, nondogmatic, and theoretically in-

[8] Maxwell and Delaney, among the graduate texts, and Rosnow and Rosenthal among the undergraduate texts, providing the most current and explicit treatment of more contemporary philosophy of science.

formed conception of science and its methods; an approach to science that is sensitive to the limitations and dilemmas of science work as well as to its political and moral implications. This approach seeks, in fact, to temper and more modestly situate the very assumptions of independence and determinacy discussed above and, as such, reflects the historicized and humanized account of science typical of much recent philosophy of science. An outcome of this sort is, of course, only possible in partnership with students and so transparency about the purposes of such contextual methods education would be essential. Students are often socialized into wanting the most straightforward, purely factual, and practically relevant information and some work would undoubtedly be necessary to engage them in so much critical and reflective work.

Conclusion

My purpose here has been to provide a faithful and fairly comprehensive account of the philosophy of science expressed in commonly used methods texts. Space prohibited any detailed critique of the science theory reflected in that analysis, but my general assessment of its positivist character and of its primary focus on middle-20th century philosophy of science accorded substantially with existing critiques. It is important, however, to note where my analysis of these texts suggested some softening in discussions of many long-dominant assumptions—in particular, those relating to experimentation, quantification, and realism. Drawing inspiration from attempts in science education to address similar inadequacies in philosophy of science coverage, I have also suggested some possibilities for reimagining psychology's methods pedagogy. More than merely updating or expanding philosophy of science content in our methods texts or classes, I have suggested that methods education in psychology could integrate postpositivist philosophies of science into a critical framework for science teaching itself.

References

Bickhard, M. H. (2001). The tragedy of operationalism. *Theory & Psychology, 11,* 35–44. http://dx.doi.org/10.1177/0959354301111002

Braaten, M., & Windschitl, M. (2011). Working toward a stronger conceptualization of scientific explanation for science education. *Science Education, 95,* 639–669. http://dx.doi.org/10.1002/sce.20449

Breen, L. J., & Darlaston-Jones, D. (2010). Moving beyond the enduring dominance of positivism in psychological research: Implications for psychology in Australia. *Australian Psychologist, 45,* 67–76. http://dx.doi.org/10.1080/00050060903127481

Burman, E. (1997). Minding the gap: Positivism, psychology, and the politics of qualitative methods. *Journal of Social Issues, 53,* 785–801. http://dx.doi.org/10.1111/j.1540-4560.1997.tb02461.x

Costa, R. E., & Shimp, C. P. (2011). Methods courses and texts in psychology: "textbook science" and "tourist brochures". *Journal of Theoretical and Philosophical Psychology, 31,* 25–43. http://dx.doi.org/10.1037/a0021575

Cozby, P. C. (2009). *Methods in behavioral research*. Boston, MA: McGraw-Hill Higher Education.

Danziger, K. (1997). *Naming the mind: How psychology found its language*. Thousand Oaks, CA: Sage.

Faulconer, J. E., & Williams, R. N. (1985). Temporality in human action: An alternative to positivism and historicism. *American Psychologist, 40,* 1179–1188. http://dx.doi.org/10.1037/0003-066X.40.11.1179

Giddings, L. S. (2006). Mixed-methods research: Positivism dressed in drag? *Journal of Research in Nursing, 11,* 195–203. http://dx.doi.org/10.1177/1744987106064635

Gravetter, F. J., & Wallnau, L. B. (2009). *Statistics for the behavioral sciences*. Belmont, CA: Wadsworth.

Green, C. D. (2001). Operationism again: What did Bridgman say? What did Bridgman need? *Theory & Psychology, 11,* 45–51. http://dx.doi.org/10.1177/0959354301111003

Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research*. Fort Worth, TX: Harcourt College Publishers.

Kuhn, T. S. (1996). *The structure of scientific revolutions* (3rd ed.). Chicago, IL: University of Chicago Press. http://dx.doi.org/10.7208/chicago/9780226458106.001.0001

Kutney, A. M. (2006). An examination of psychiatric-mental health outcomes from the perspectives of logical positivism and phenomenology. *Journal of the American Psychiatric Nurses Association, 12,* 22–27. http://dx.doi.org/10.1177/1078390306286443

Laungani, P. (2002). Cross-cultural psychology: A handmaiden to mainstream Western psychology. *Counselling Psychology Quarterly, 15,* 385–397. http://dx.doi.org/10.1080/0951507031000069392

Leiviskä, A. (2013). Finitude, fallibilism and education towards non-dogmatism: Gadamer's hermeneutics in science education. *Educational Philosophy and Theory, 45,* 516–530. http://dx.doi.org/10.1080/00131857.2012.732012

Luyten, P., Blatt, S. J., & Corveleyn, J. (2006). Minding the gap between positivism and hermeneutics in psychoanalytic research. *Journal of the American Psychoanalytic Association, 54,* 571–610. http://dx.doi.org/10.1177/00030651060540021301

Lynd-Stevenson, R. M. (2007). Concerns regarding the traditional paradigm for causal research: The unified paradigm and causal research in scientific psychology. *Review of General Psychology, 11,* 286–304. http://dx.doi.org/10.1037/1089-2680.11.3.286

Martin, D. W. (2008). *Doing psychology experiments*. Belmont, CA: Thomson/Wadsworth.

Martin, J. (2003). Positivism, quantification and the phenomena of psychology. *Theory & Psychology, 13,* 33–38.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective*. Mahwah, NJ: Erlbaum.

Michell, J. (2003). The quantitative imperative: Positivism, naïve realism and the place of qualitative methods in psychology. *Theory & Psychology, 13,* 5–31. http://dx.doi.org/10.1177/0959354303013001758

Niaz, M. (2010). Science curriculum and teacher education: The role of presuppositions, contradictions, controversies and speculations vs Kuhn's 'normal science'. *Teaching and Teacher Education, 26,* 891–899. http://dx.doi.org/10.1016/j.tate.2009.10.028

Pearce, J. V. (2013). The potential of perspectivism for science education. *Educational Philosophy and Theory, 45,* 531–545. http://dx.doi.org/10.1080/00131857.2012.732013

Pelham, B. W., & Blanton, H. (2007). *Conducting research in psychology: Measuring the weight of smoke*. Belmont, CA: Thomson Wadsworth.

Proctor, R. W., & Capaldi, E. J. (2001). Improving the science education of psychology students: Better teaching of methodology. *Teaching of Psychology, 28,* 173–181. http://dx.doi.org/10.1207/S15328023TOP2803_02

Rodríguez, M., & Niaz, M. (2004). The oil drop experiment: An illustration of scientific research methodology and its implications for physics textbooks. *Instructional Science, 32,* 357–386. http://dx.doi.org/10.1023/B:TRUC.0000044641.19894.ed

Rosnow, R. L., & Rosenthal, R. (2008). *Beginning behavioral research: A conceptual primer*. Upper Saddle River, NJ: Pearson/Prentice Hall.

Santos, W. D. (2009). Scientific literacy: A Freirean perspective as a radical view of humanistic science education. *Science Education, 93,* 361–382. http://dx.doi.org/10.1002/sce.20301

Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2009). *Research methods in psychology*. Boston, MA: McGraw-Hill Higher Education.

Sommer, R., & Sommer, B. B. (2002). *A practical guide to behavioral research: Tools and techniques*. New York, NY: Oxford University Press.

Tissaw, M. A., & Osbeck, L. M. (2007). On critical engagement with the mainstream. *Theory & Psychology, 17,* 155–168. http://dx.doi.org/10.1177/0959354307075120

Wagner, C., & Maree, D. (2007). Teaching research methodology: Implications for psychology on the road ahead. *South African Journal of Psychology, 37,* 121–134. http://dx.doi.org/10.1177/008124630703700109

Wertz, F. J. (1999). Multiple methods in psychology: Epistemological grounding and the possibility of unity. *Journal of Theoretical and Philosophical Psychology, 19,* 131–166. http://dx.doi.org/10.1037/h0091173

Zembylas, M. (2000). Something "paralogical" under the sun: Lyotard's postmodern conditionand science education. *Educational Philosophy & Theory, 32,* 159–184.

Zembylas, M. (2006). Science Education as Emancipatory: The case of Roy Bhaskar's philosophy of meta-Reality. *Educational Philosophy and Theory, 38,* 665–676. http://dx.doi.org/10.1111/j.1469-5812.2006.00218.x

Received July 21, 2015
Revision received March 15, 2016
Accepted March 16, 2016 ■