The Journal of Mind and Behavior
Spring 2004, Volume 25, Number 2
Pages 123–144
ISSN 0271–0137

Why Psychology Hasn't Kept Its Promises

Henry D. Schlinger

California State University, Northridge and *University of California, Los Angeles*

This essay posits that psychology's general lack of respect as a science stems from two related problems: the continued focus on conceptually vague mentalistic constructs and the adherence to a methodology that emphasizes statistical inference over experimental analysis. The lack of a thoroughgoing experimental analysis has so far prevented psychologists from discovering a set of foundational principles thus inhibiting them from being able to predict and control individual behavior. Psychologists can remake their conceptual and methodological foundations by focusing on the relationship between observed behavior and its context and by adopting methods of experimentation that would aid in the discovery of orderly functional relationships. This knowledge could be used to more parsimoniously explain complex behavior, including cognitive phenomena, and to more effectively solve practical problems.

Keywords: scientific methodology, deductive inference, behaviorism

> The woods are lovely, dark, and deep,
> But I have promises to keep,
> And miles to go before I sleep,
> And miles to go before I sleep
>
> Robert Frost

In *How to Think Straight About Psychology* (1998), Stanovich paints a picture of psychology, in the first chapter, as a healthy science — "Psychology is Alive and Well and Doing Fine Among the Sciences" — and in the final chapter as a field with an image problem — "The Rodney Dangerfield of the Sciences." If psychology is indeed a healthy science, then why does it "get no respect"? Contrary to Stanovich's answer that psychology is plagued by mis-

The author is grateful to Jay Moore, Ed Morris, David Palmer, Julie Riggott, Ray Russ and Stuart Vyse for their insightful comments on earlier drafts of this essay. Requests for reprints should be sent to Henry D. Schlinger, Ph.D., Department of Psychology, California State University, Northridge, 18111 Nordhoff Street, Northridge, California 91330–8255. Email: hschling@ csun.edu

understanding and misrepresentation, others don't think psychology is doing so well (e.g., Holth, 2001; Loftus, 1996; Lykken, 1991; Machado, Lourenço, and Silva, 2000; Slife and Williams, 1997; Staats, 1983). For example, Machado et al. (2000) argue that psychology "has a credibility problem because it has promised more than it has delivered" (p. 20) and they cite others who conclude that if science is defined by its accomplishments rather than by its method then psychology is not a science.

In the present article, I go further and contend that psychology has produced very few noteworthy discoveries: it has offered few if any satisfactory explanatory concepts and, therefore, has not advanced in the same way as the other sciences. I argue that the main culprit is psychologists' continued emphasis on mind, and though nowadays it is dressed up in the fancier metaphorical garb of information processing or neuralese, it is the same mind about which pre-scientific philosophers speculated. In short, mentalism stands in the way of psychology joining the ranks of the other sciences.

The emphasis on mind instead of behavior is related to another culprit. Unlike the natural experimental sciences (e.g., functional biology, chemistry), in which scientists carry out experimental analyses involving the systematic manipulation of precisely defined and measured independent variables, psychology is wedded to a methodology that relies on hypothesis testing and inferential statistics that, according to some is relatively bankrupt (e.g., Loftus, 1991; Machado et al., 2000). I argue that this methodological shortcoming has contributed to psychology's failure to become a truly experimental science, precluding the scientific goals of prediction and control. These two failings combine to be a formidable obstacle to psychology's membership in the club of the natural sciences and constitute the main reasons why, in my view, psychology gets no respect as a science.

Delivering the Goods

To understand the roots of psychology's credibility problem, let us revisit two purposes of science. Palmer (1998) explains:

> Science serves us in two ways. First, it underpins our mastery of the physical and biological world: We should not like to do without our vaccines, antibiotics, semiconductors, or internal combustion engines. But perhaps an even more important service is to resolve mysteries about nature. Science offers beautiful, elegant and often, deeply satisfying explanations for complexity and order in nature, and if forced to choose, we might prefer to live in a cave, with our understanding intact, than in a wonderland of gadgets, benighted by superstition. (p. 3)

To paraphrase Schwartz (1989), simply, psychology has not "delivered the goods." This has occurred because much of psychology lacks critical features that characterize the experimental sciences, including a solid foundation of experimentally derived principles and a sound conceptual framework tied to those principles.

As a result, psychology also lacks cohesion among many of its various subspecialties. Unlike the biological sciences, for example, that are united by common principles and overarching theories (e.g., natural selection theory, cell theory, etc.), psychology is an assortment of sub-disciplines that are not unified by any common principles or basic units of analysis.[1] For example, the areas of social, developmental, personality, and cognitive psychology all seem to have different subject matters, each populated with different theories or models. Except for an overall cognitive orientation that pervades much of psychology nowadays, each of these areas is relatively independent. Some authors have argued that psychology's disunity is not necessarily a bad thing (Bower, 1993). Others have contended that it may be no different than the other sciences, for example, biology (Viney, 1996). However, as Machado et al. (2000) have noted, "the fragmentation of the field and the artificial specialization of its members should not be confused with the division of labor that exists in sciences such as physics and biology" (p. 12).

Most importantly, though, psychology has no universally accepted, objectively defined subject matter (Is it the scientific study of mind or behavior, or both? And if it is mind, what is mind and how can that be studied scientifically?), and consequently, no commonly accepted basic units of analysis. As Palmer (1991) explains, "It is important to identify the unit of analysis in any scientific endeavor, but it is especially important in a domain with a rich prescientific vocabulary" (p. 262). Psychologists' "prescientific vocabulary," an inheritance from their philosophical ancestors, includes such terms as *mind*, *memory*, *thinking*, and *consciousness*, terms psychologists and philosophers have struggled with (unsuccessfully, I believe) for centuries.

One way to evaluate psychology's promises is to determine how well it has realized the purposes of science. In introducing psychology as a science, some textbooks describe several goals of science from which we can distill a list of three: objective observation and measurement, control through experimentation, and prediction (e.g., Bernstein, Clarke–Stewart, Penner, Roy, and Wickens, 2000; Carlson and Buskist, 1997; Huffman, 2002). Meeting these goals would help psychologists to offer satisfactory explanations of phenomena comprising their subject matter and to enhance their mastery of nature (i.e., to change the behavior of individuals). Psychology's failure to develop

[1]Although they offer different perspectives, numerous authors have noted the fragmentation or disunity of psychology (e.g., Bower, 1993; Koch, 1993; Slife and Williams, 1997; Staats, 1991, 1999; Viney, 1996).

as a science that can offer accurate prediction and control of behavior and elegant explanations can be traced directly to its continued emphasis on mental events.

The Promise of Objective Observation and Measurement

Objective observation is the foundation of science and that feature which distinguishes it from other attempts to understand nature such as philosophy and theology. On the surface, this seems an acceptable assertion until we realize that the very subject matter of psychology — mind — fails to meet this goal. It is true that just because something can't be observed, doesn't mean it doesn't exist; but, at the very least unobservability does pose serious methodological problems. Many psychologists are quick to point out that although mental or cognitive events can't be directly observed, the behavior that reflects them can be. As the author of a prominent textbook in cognitive psychology puts it: "Although things on the surface, the behaviors, are important elements, there appear to be 'internal' systems which are behind such actions . . ." (Solso, 2001, p. 4). [And he is not talking about underlying neurophysiology.] This locution seems perfectly reasonable to most psychologists and lay people. For example, the behavior of speaking is said to reflect ideas, thoughts, meanings, and intentions or what some have called "mentalese" (e.g., Pinker, 1994). To inform them about the underlying ideas, thoughts and meanings, psychologists observe speaking. Psychologists don't fail the goal of objective observation completely because they must, by necessity, observe and measure overt behavior. But those who call themselves cognitive psychologists do not and cannot directly observe or measure the real dependent variables of interest — cognitions — because that is not possible. Does the rationale for observing behavior solely to infer cognitive events hold up to scientific and logical scrutiny?

The Illusion of Cognition

The questions about cognitions are simple. What are they? Where are they? Can they be directly measured? Is it possible to explain observed behavior without appealing to cognition? Consider memory. Most psychologists use the term *memory* to refer to a collection of cognitive processes (e.g., encoding, storage and retrieval) that is divided into different storage components — sensory, working, and long-term memory — and classified as implicit (including procedural) and explicit (including declarative, episodic, and semantic), among others. But exactly what processes and structures make up memory? Are the structures and processes homologous with central nervous system structures and processes or are the memory processes produced by the nervous system as some cognitive psychologists assert (e.g., Solso, 2001)? If the latter, then how can we objectively observe and study memory as a datum separate from the behavior said to reflect it?

By cognitive psychologists' own admission, the only evidence of memory is behavior in certain contexts; that is, cognitive psychologists infer the structures and processes of memory solely from overt behavior (Baars, 1986; Solso, 2001). This can potentially lead to some less than logical ways of talking about memory. For example, if I see a friend and utter his name, it is said that I *remembered* his name. As long as remembering isn't asserted to be anything more than saying my friend's name when I see him — both observed and potentially alterable variables — then there is no problem. Often, however, it is said that I have (or possess) a visual (called *iconic*) memory of my friend. This locution implies that the memory is separate from the actual behavior of saying the name (i.e., it is reified), a position that has also been criticized as being essentialistic in that the memory is viewed as a thing (e.g., a representation that is stored) [Palmer and Donahoe, 1992]. In addition, asserting a memory process separate from the behavior of saying the name implies that the memory in some way causes the behavior, a distinctly mechanistic account (Morris, 2003). Finally, appealing to the encoded memory as an explanation for the observed behavior of saying my friend's name is circular (or tautological) if the only evidence for the encoding and long-term memory of my friend's name is my saying it when I see him.

Claiming that memory, or any other cognitive process, is real, independent of observed behavioral or neurological functions said to reflect it, is not only essentialistic but it leaves the ghost of dualism to haunt modern psychology. This Cartesian quandary was no more clearly evidenced than when a noted neuroscientist asked, "How does the nonmaterial mind influence the brain, and vice versa"? (Fischbach, 1992, p. 48). Even a popular textbook on cognitive psychology opens a chapter with the following dualistic assertion: "Remarkably, we humans occupy two worlds at the same time" (Solso, 2001, p. 36). And although the author asserts that there is no fundamental difference between the two worlds, by speaking of mind *and* brain (e.g., "This chapter is about both mind and brain . . ." p. 38), he leaves the Cartesian door wide open. Or when he states that "[A]ll cognition is the result of neurological activity" (p. 38), one is compelled to ask what the nature of cognition is if it is not behavior.

One thing is certain, psychologists have never directly observed or measured memory (or any other cognitive event or process) conceived in this manner, which leads to descriptions that are couched in metaphorical terms such as *cognitive maps*, *engrams*, *encoding* and *retrieval*, *sensory registers*, and *storage*. Palmer (2003) notes:

> Unfortunately each such concept introduces a qualitatively new element that itself requires explanation or justification. That is, each new term must ultimately be paid for in the coin of physical, biological, or behavioral events. Like the improvident debtor who pays off one credit card by drawing down the balance of a second, such devices provide only temporary satisfaction, for the overall explanatory burden has been increased, not reduced. (p. 172)

Another problem resulting from an emphasis on cognitive events is that in the absence of direct observation presumably different cognitive processes become indistinguishable. For example, when I utter my friend's name, it may be said that I *remember* his name, that I *perceive*, or *recognize* him, or that I have some *understanding* or *knowledge* or *representation* of him. In this instance, memory, perception, recognition, understanding, knowledge and representation are presumed to be different cognitive processes, but the evidence for all of them is the same — saying my friend's name when I see him. And neurophysiological data, even those from the currently fashionable neuroimaging methods, don't support the distinctions because it is not now possible to independently observe the corresponding brain functions and correctly predict the behaviors said to reflect memory, perception, recognition, or understanding.

Memory and other mental concepts, although acceptable in casual discourse as non-technical labels for behavior, do not map easily onto experimentally established neurophysiological (or behavioral) functional units of analysis (see below) and, thus, remain illusory. Moreover, many of the behaviors said to reflect memory, perception and other cognitive processes can often be explained more parsimoniously with experimentally established general principles of behavior (e.g., Palmer, 1991, 2003; Schlinger, 1992, 1995, Chapter 4).

Deductive Inference in Psychology: The Cart Before the Horse

Some may answer that inferring unobserved cognitive processes based on observed behavior is similar to the practice in the natural sciences of regularly hypothesizing unobserved events. Let us examine this claim briefly in the context of an analogy suggested by Schlinger (1998). Over the last couple of decades or so, astronomers have debated the existence of extrasolar planets. Until recently, astronomers could not observe the planets directly because the light from the stars around which the planets orbit obscures the planets themselves. Instead, astronomers inferred (or deduced) the existence of these planets from indirect observations of the slight wobbles in the star's orbit caused by the gravitational pull of the presumed planets. In fact, such deductions of unobserved phenomena are not uncommon in the mature sciences. Some may argue, then, that inferring unobserved cognitive events based on observations of behavior has scientific precedent. There is a difference, however. Astronomers knew about the gravitational effects one body exerts on another from experiments carried out with observed bodies here on earth. Once scientists began to understand the effects of gravity from direct experiments on observed bodies, they could then extrapolate those findings to other bodies that couldn't be manipulated. They assumed, however, that the bodies differed only in location and scale, not in type. For example, Newton reasoned that the moon exerted a gravitational pull on the ocean tides because he had experimented with things closer and smaller. Psychologists, on the other hand, measure events of one type — behavior — in order to talk about events of a completely different type — cognitions. Finally, astronomers intuitively adhere to the principle of parsimony, which states that the explanation of an observation requiring the fewest assumptions is to be preferred, whereas psychologists' theorizing is often based on assumptions (about the existence of cognitive structures and processes) that can never be directly tested. In most cases, inferring cognitive events and processes to explain behavior is not the most parsimonious approach because there are alternative interpretations based on experimentally derived principles (Palmer, 2003; Schlinger, 1993).

Rather than deriving from direct observation, cognitive constructs, such as *memory*, are posited based on the combination of misleading subjective introspection, on idiosyncrasies of our shared commonsense psychological vocabulary, and on the complex and often subtle nature of the real causes of behavior, the discovery of which is the ultimate challenge for psychologists. As psychologists themselves have noted, human beings often think irrationally about cause and effect relationships (e.g., Gilovich, 1991; Vyse, 1997). Obviously, psychologists are not immune to these logical stumbles.

In concentrating on unobserved events before first trying to understand observed events, cognitively oriented psychologists have put the cart before the horse. The alternative is to attempt to understand observed behavior in its context before speculating about unobserved phenomena. In order to explain why we say the names of other people when we see them, a more productive scientific approach than speculating about cognitive constructs would be to conduct an experimental analysis of behavior of saying names (e.g., of objects or individuals). Once orderly relationships are discovered between objective events and behavior (e.g., seeing or hearing a person or an object and saying its name), scientists are in a better position to predict and control the behavior. This tactic doesn't negate the possibility of unobserved events as either proximate causes of behavior or as behaviors themselves; rather, it emphasizes a more conservative approach that resembles the historical development of the other sciences in which observed events are studied and understood before contemplating unobserved events. Once observed events are understood, scientists can then postulate that unobserved events obey the same laws. As Palmer (2003) notes, "We can consider observability to be a continuum, and since the boundary between the overt and covert is an arbitrary and variable point along that continuum, we can assume that the laws of observed behavior hold for the unobserved" (p. 178). This position is at odds with much of contemporary cognitive psychology, which assumes that separate laws and rules exist for unobserved cognitive processes.

Once scientists observe (and measure) some phenomenon, they are in a better position to describe it in terms that can aid other researchers in making the same observations. Again, mental or cognitive events fail this standard. Such descriptions, when they are offered, are, by necessity, in the form of metaphor. Consider the following brief excerpt from a recent article on the tip-of-the-tongue (TOT) phenomenon:

> Results support the transmission deficit model that the weak connections among phonological representations that cause TOTs are strengthened by production of phonologically related words. (James and Burke, 2000, p. 1378)

This short quotation is almost entirely metaphorical. In other words, very little that is described is observed — neither the representations nor the (weak or strong) connections between them — despite the fact that the description is based solely on the performance of participants in an experiment. Worse, the weak connections among the representations are said to be the cause of the TOTs. The exact structures (phonological representations) and processes (strengthening of connections by production of phonologically related words) involved are never specified, thus making it impossible, in the absence of descriptions of actual behavior, for other researchers to make the same observations.

Speaking metaphorically is not bad per se; it does make difficult concepts more accessible to the uninitiated. However, as Machado et al. (2000) write, "when a scientist extends an everyday concept into a new domain, it is incumbent upon him to explain the new meaning of the concept and how the concept should be used" (p. 33). Consider the language acquisition device (LAD) metaphor. Neither psycholinguists nor psychologists can describe, much less explain, a LAD in terms of any known scientifically derived principles from psychology or biology (Palmer, 2000). The problem for cognitive psychology is that the entire field "is metaphorical in character" (Solso, 2001, p. 22). This is because the dependent variables of interest — cognitions — are never observed. Psychologists are forced to describe cognitive events and processes in information-processing terms (e.g., encoding, retrieval, central executive). But the problem goes even further. Even such seemingly benign concepts as *memory, perception, thinking, consciousness*, as well as a variety of other standard mentalistic terms cannot be described separately from behavior with any terms that can lead to their independent observation, measurement, description, or interconnectedness with other objectively defined events (e.g., neurophysiological structures or processes).

Lacking accurate descriptions of the basic datum of a discipline obviously poses problems for an aspiring science, not the least of which involves controlling the subject matter through experimentation. Ironically, because of the intangible and unobserved nature of cognitive phenomena, psychologists are forced to always operationalize them as observed behavior, but not because they view observed behavior as an important subject matter in its own right. Even so, psychologists' attempts at experimentation often fall short of those necessary to provide elegant, parsimonious ultimate explanations of the phenomena of interest.

The Promise of Experimentation

Psychologists talk about the importance and power of experimentation, but how well do their actions match their words? In fact, most psychological experiments are not capable of uncovering orderly functional relationships between independent and dependent variables because they do not demonstrate unambiguous experimental control over the dependent variables. The reasons are many. For one, most experiments are carried out with groups of subjects instead of isolating the controlling variables of the behavior of one individual. Second, the independent variables are often too complex or ambiguous or they are defined in such as way that their critical values cannot be isolated. For example, a recent experiment by Bushman and Bonacci (2002) compared recognition and recall of several advertised products among 324 participants divided into three groups who watched either a violent, sexually explicit, or neutral TV program that contained nine ads. In this experiment the independent variables — the three types of content of the shows and the nine ads — were hardly specific and well defined. Third, sometimes the research designs are dictated largely by the statistical analysis that will be used to interpret the results. Fourth, many psychological experiments still incorporate data from surveys, self-reports, or other measures, which although not completely without merit, are notoriously unreliable because of their subjectivity. These problems occur, in part, because with few exceptions psychology has no basic units of analysis and, in part, because psychologists follow a rigid recipe taught to all psychology students in research methods courses rather than conducting experimental analyses.

This was not always the case, however. After John Watson (with his 1913 manifesto) initiated what some have called a revolution in psychology in which behaviorism supplanted (or overthrew) the existing psychology of consciousness, researchers began to experimentally analyze the relationship between environmental variables and behavior. From thousands of such experiments in the ensuing almost 100 years, psychologists were able to discover functional relationships that displayed enough consistencies as to be formulated into general principles. Many of these researchers were called *behaviorists* and some, including those who also subscribed to Skinner's philosophy of radical behaviorism, began to extend the basic principles gleaned from experimentation with non humans and humans to more complex behaviors in humans, such as remembering, perceiving, talking, and thinking. But apparently these efforts were either too slow or too unsatisfying for other psychologists.

Some scholars now claim that a second revolution occurred in the 1950s and 1960s in which cognitivism overthrew behaviorism as the dominant system in psychology (e.g., Baars, 1986; Gardner, 1985). Although not every historian of psychology subscribes to the claim of a cognitive revolution (e.g., Leahey, 1992), and although there are still thousands of practicing behaviorists and learning researchers all over the world, there is no doubt that the prevailing metatheory in psychology is a cognitive one. Some have argued that contrary to a cognitive revolution, a form of behaviorism, called *mediational* (or *methodological*) *behaviorism*, simply evolved to overtake Skinner's radical behaviorism with which it co-exited for several decades (Leahey, 1992; Moore, 1995). As Morris (2003) points out, "Cognitivism's attempt to remake mentalistic Western folk psychology on the pattern of the physical sciences has proved to be a daunting task" (p. 276). We may go further and question how well cognitive psychologists have succeeded in creating either satisfying explanations of behavior or a technology for solving practical problems. As I have argued in this essay, the reasons for these failures are psychologists' emphasis on unobserved mental constructs and the ways in which they carry out experiments. These differences may have prevented psychological researchers from discovering a set of universal foundational principles.

Units of (Experimental) Analysis

The key to the emergence of laws and principles is the discovery of functional units of analysis. The success of any science, thus, depends on whether that science has discovered "proper units" of analysis. As Zeiler (1986) has noted,

> A well-defined unit clarified the way phenomena are conceptualized and thereby guides research and theory. Isolation of a unit brings order to otherwise discrepant phenomena; invalid units easily lead to confusion as to the meaning and significance of the data. (p. 1)

The discovery of the cell, as a basic unit of biology, integrated the disciplines of anatomy, embryology, botany, and zoology into the field that we now call the biological sciences (Zeiler, 1986). With few exceptions (e.g., behavior analysis) psychology has no such analytic units (see Glenn, Ellis, and Greenspoon, 1992; Schlinger, 1992; Zeiler, 1986). This is perhaps why many of the various sub-disciplines in psychology seem so disparate. In particular, cognitive psychology has "no body of systematized principles, no unique set of data, no characteristic measurement techniques, and no typical investigative procedures" (Sidman, 1986, p. 214). One reason for these shortcomings is that the basic units of cognition, such as representations and memories, are not functional units that can be subjected to systematic experimental analysis. Rather, such cognitive concepts are really complex behavioral events that are, themselves, compounds of more elementary behavior-analytic units (Palmer, 2003). As Palmer has noted, because of "the complexity of cognitive phenomena and the difficulty of collecting reliable data, any account of cognition must be tentative and cautious, a circumstance that has encouraged a profusion of competing models and theories" (2003, p. 168). This situation makes cognitive concepts ripe for an interpretation using other, more established, principles. For example, perhaps interpretations of complex human behavior based on decades of experimentally established behavior analytic units should be exhausted before constructing broad inferential metaphorical models. It is, thus, noteworthy that Skinner, accused by numerous psychologists of disregarding or ignoring private events, actually defined his radical behaviorism by its consideration of private events and offered it as an alternative to cognitive psychology:

> A science of behavior must consider the place of private stimuli as physical things, and in doing so it provides an alternative account of mental life. The question, then, is this: What is inside the skin, and how do we know about it? The answer is, I believe, the heart of radical behaviorism. (1974, p. 233)

A Recipe for Research

Basic experimental analysis, once more common in psychology, has been largely replaced by a formal set of methods called the hypothetico–deductive approach. One might say that Hullian behaviorism won out over Skinnerian behaviorism is this regard. In the hypothetico–deductive approach, observations are made and then hypotheses about certain relationships are predicted or deduced, usually by comparing groups of subjects. Research is designed to confirm (or refute) the hypotheses, which ultimately strengthens or weakens a theory.

As a result, very few experiments manipulate precisely defined independent variables and look for order in the relationship between them and the dependent variables the way that chemists and experimental biologists do. Moreover it is not uncommon for each study to propose a different model or theory of the specific phenomenon under investigation. The implication is that each psychological phenomenon (or behavioral oddity) requires its own explanation or theory, an approach, which is at odds with the natural sciences in which fewer general and elegant theories encompass many often seemingly unrelated phenomena. As Hempel (1966) explains, a good theory

> offers a systematically unified account of quite diverse phenomena. It traces all of them back to the same underlying processes and presents the various empirical uniformities they exhibit as manifestations of one common set of basic laws. (p. 75)

With few exceptions, psychology offers no theory that accomplishes what a good theory must. One reason may be the almost universal reliance on group designs.

The More the Muddier

It seems axiomatic in psychology that research must involve groups of subjects. It is common knowledge that experiments employ one or more experimental and control groups. But group designs come at a price. First, there are differences between individuals in the group even if the researchers truly select participants randomly and match for gender, age, and so on. Such variability originates mostly from different learning experiences, but also from genetic differences. In an attempt to control for this variability, researchers are forced to design experiments according to how the resulting data will be analyzed statistically. As Machado et al. (2000) have stated, there is a troubling tendency in psychology "to assess the significance of data by the means used to obtain them" (p. 6). [The alternative is to compare many different values of an independent variable on the behavior of individuals which is much more time-consuming and effortful.] The price often paid is that such experiments do not permit the researchers to predict and control the behavior of individuals. As Skinner noted, "No one goes to the circus to see the average dog jump through a hoop significantly oftener than untrained dogs . . ." (1956, p. 228).

The Formal vs. Informal Approach to Research

As early as 1956, Skinner lamented that the formalized methods practiced by psychologists had little to do with the actual behavior of scientists who have made significant discoveries. He claimed that such methods had acquired a "purely honorific status" in psychology. Almost fifty years later, Machado et al. (2000) worried that the "standard conception" of the research method among psychologists "uses powerful statistical techniques but has rarely delivered equally powerful concepts, functions, and theories" (p. 6).

The American Psychological Association even convened a special task force to address growing concerns about the domination of statistical inference testing methods in psychological research (Wilkinson and Task Force on Statistical Inference, 1999), an effort spurred by questions raised by other psychologists (e.g., Cohen, 1994; Loftus, 1991, 1996). Consider Loftus' (1991) appraisal of hypothesis testing in psychology:

> First, for mid-20th-century experimental psychologists (status hungry, perhaps, amidst their natural-science colleagues) hypothesis testing provided the illusion of endowing psychological data — which are intrinsically complicated, messy, multidimensional, and subjective — with a seductive simplicity and objectivity. Banished was the slovenly panoply of "eyeballing curves, personal judgment, description without inference, or bargaining with the reader" to be replaced by one clean, simple, refreshing rule: if $p < .05$ (or so) an effect is real; otherwise, it's not. In the domain of pure research, this switch eased the decision process for journal editors, while in the domain of practical applications (particularly educational and military applications) it provided researchers with a new concept, easily explainable to policy makers, that could be used to justify (or denounce) the implementation of novel techniques. (p. 104)

So what are psychologists doing wrong? The present essay urges a stronger solution than Loftus' emphasis on a return to the use of simpler tools such as probability theory, measurement, and descriptive statistics: a return to the inductive science practiced by such pioneers as Pavlov, Skinner, and their descendants in which the discovery of sensitive dependent variables (e.g., rate of response) aided in the discovery of functional units of analysis (e.g., operants, respondents) and general principles.

The hypothetico–deductive (hypothesis-testing) approach practiced by most psychologists differs from its cousin in the mature laboratory sciences. In the mature sciences, the approach begins more inductively. As Bertrand Russell (1961) noted, "Scientific method seeks to reach principles inductively from observed facts" (p. 58). In the mature experimental sciences, orderly relationships between variables are derived through systematic experimental manipulation, and any repeatable regularities that are discovered are summarized verbally or mathematically as general principles or laws. According to Russell, "The test of scientific truth is patient collection of facts, combined with bold guessing as to laws binding the facts together" (p. 514), and "The thing that is achieved by the theoretical organization of science is the collection of all subordinate inductions into a few that are very comprehensive — perhaps only one" (p. 530). These inductions, or principles, become the theoretical bedrock of the discipline from which hypotheses about novel relationships can then be deduced. Skinner (1956) believed that this more informal approach also increases the probability of the kinds of accidental, serendipitous, discoveries that have characterized the mature sciences for centuries.

In an early autobiographical account of his own experimental activity that led to the elucidation of the principles of operant learning, Skinner (1956) described in a kind of tongue-in-cheek fashion several informal principles of science, for example, "When you run onto something interesting, drop everything else and study it" (p. 223). This attitude, along with two of Skinner's other informal principles — "Some people are lucky" (p. 225) and "Apparatuses sometimes break" (p. 225) — led not only to Skinner's own discoveries of some of the methods and principles of operant learning, but to other scientific findings in the history of science, such as Pavlov's discovery of conditioned reflexes and Alexander Fleming's discovery of penicillin, to name but a few.

Because psychologists generally follow a rigid set of rules for doing research, or as Machado et al. (2000) call it, "the standard conception" of the scientific method, they are not likely to notice interesting deviations from their expected results, much less to abandon their statistically driven experimental designs to study them. If equipment breaks, researchers are likely to fix it and resume their study and any unintended data are probably tossed out. And luck is just another name for what Pasteur meant when he said that, "chance favors the prepared mind." Unlike Skinner, Pavlov, Fleming, Pasteur and many other scientists, psychologists are generally less prepared for chance occurrences because they are not looking for order in their subject matter, but, rather, order in their statistical analyses.

In contrasting his research activity, and by implication, that of other scientists, with that of the more formalized approach followed by many psychologists, Skinner (1956) said,

> I never faced a Problem which was more than the eternal problem of finding order. I never attacked a problem by constructing a Hypothesis. I never deduced Theorems or submitted them to Experimental Check. So far as I can see, I had no preconceived Model of behavior — certainly not a physiological or mentalistic one and, I believe, not a conceptual one . . . . Of course, I was working on a basic Assumption — that there was order in behavior. (p. 227)

Skinner's view on theory is similar to that of Russell's. Chiesa (1992) explains:

> *Theory* may also be used to refer to an explanatory system, such as Skinner's, that describes regularities, states general principles, and integrates uniformities in a given subject matter. These . . . kinds of theories do not carry the same requirement to be submitted to experimental check because they are "data driven" (derived from observation) and are not constructed prior to experimentation. (p. 1294)

It is not wrong for psychology to emulate the physical sciences in adopting the hypothetico–deductive approach; it is simply putting the cart before the horse in the development of a truly scientific psychology. A science develops when the subject matter is objectively delineated and that happens when researchers begin their search for order among the independent and dependent variables of the science.

Demonstration vs. Explanation: Three Illustrative Examples

I have argued that most experiments in psychology do not achieve the goal of experimentally controlling the subject matter such that orderliness between independent and dependent variables can emerge. Instead they are at best demonstration experiments. That is, they succeed in demonstrating some of the conditions under which some behaviors occur and even then, only statistically across groups of subjects. These conditions may comprise some of the proximate causes of behavior (see Alessi, 1992) — in other words, those factors that are the immediate precursors of the behavior. In what follows I offer examples of demonstration experiments from three content areas in psychology.

Social psychology. Social psychologists have demonstrated that various factors influence compliance. One of these factors is called the "foot-in-the-door" technique in which getting compliance with a small request (foot-in-the-door) enhances the later chances of compliance with a larger request. A related factor is the door-in-the-face technique, in which requesting a large commitment enhances later compliance with a request for a smaller commitment (e.g., Cann, Sherman, and Elkes, 1975; Freedman and Fraser, 1966).

The first problem with studies on these factors that influence compliance is that the methods always involve group designs, which means that there are instances in which getting compliance with a small request has no effect on later chances of compliance with a larger request. These individual cases are lost in the statistical analysis and are never explained. In cases where the general claim does hold true there is another problem: although the research does a reasonable job of describing behaviors we may label *compliance* by suggesting some of the circumstances under which they occur (e.g., first getting compliance to a small request), it does not explain why compliance occurs to either the small or larger request. Such an explanation would point to ultimate causes located in the learning history of individuals. In order to answer the question of why first making a small request enhances the chances of compliance with a bigger request, different values of suspected independent variables must be systematically manipulated to control the relevant behaviors such that they can be accurately predicted in a given individual. In the absence of an experimental analysis, we are left with research showing statistical relationships, but no ultimate explanations of the individual behaviors that make up those relationships.

Developmental psychology. Demonstration experiments are also common in the field of developmental psychology largely because the goal of many researchers is simply to demonstrate age-related abilities. For example, consider the extensive literature on infant–mother attachment using the "strange situation" (see Ainsworth, Blehar, Waters, and Wall, 1978), in which a child and mother enter a room, the mother leaves, a stranger enters, the mother returns, and so on. In general, research demonstrates that some children (called *securely attached*) will not cry when the mother leaves or when a stranger enters, but rather, will stay by the mother briefly, or explore or continue to play, whereas other children (called *anxiously-attached*) will cry and cling when the mother leaves or when a stranger enters. While such research demonstrates some of the immediate precursors of various attachment behaviors in children under contrived situations, it does not explain why the behaviors occur under these circumstances or why they do for some children but not for others. There is simply no experimental analysis of all of the suspected controlling variables of the behavior of one individual child. As with compliance behaviors, the ultimate causes of attachment behaviors are most likely located in the learning history of individuals (see Gewirtz and Pelaez–Nogueras, 1991, 1992; Schlinger, 1995, Chapter 9). In much of the literature on infant–mother attachment, however, the ultimate explanations of the behaviors are theorized about using mentalistic concepts such as *expectations*, *memories*, or *internal working models*. (See Fagen, 1993 and Schlinger, 1993 for an exchange on the issue of inferring expectations after operant learning in infants.)

Cognitive psychology. The study of cognitive processes such as memory and attention also reveals a high proportion of demonstration experiments. Consider the Stroop effect in which "naming the color of an incompatible color word (RED printed in green ink; say, 'green') is much slower and more error-prone than is naming the color of a control item (e.g. XXX or CAT printed in green; say 'green')" [MacLeod and MacDonald, 2000, p. 383]. Some researchers consider the Stroop effect so important, they state that, "This seemingly simple interference phenomenon has long provided a fertile testing ground for theories of the cognitive and neural components of selective attention" (MacLeod and MacDonald, 2000, p. 383). Specifically the Stroop effect is said to provide "a theoretical window on how we deal with conflicting stimuli, and it is a fertile testing ground for ideas about" the role of learning in the development of automaticity (MacLeod, 1998, p. 201), even though no principle of learning is ever mentioned.

Rather than offering ultimate explanations of the longer (or shorter) latencies or increased (or decreased) errors, or some other measure of response strength to colored words, the experiments only demonstrate the phenomenon and some of the correlated conditions under which it occurs. Recently, researchers using neuroimaging techniques have begun to speculate about the neural underpinnings of the phenomenon (e.g., Banich et al., 2000) and others have suggested connectionist models to try to explain it (e.g., Cohen, Dunbar, and McClelland. 1990; MacLeod and Macdonald, 2000). However, notwithstanding the methodological and interpretive problems inherent in neuroimaging studies (see Faux, 2002), identifying the parts of the brain that become active during different tasks at best only clarifies the structures that are active when the behaviors are observed — it doesn't explain the behaviors. And, despite the appeal of a connectionist model, such a model does not clarify the ultimate causes (e.g., a rich reinforcement history for reading words that are the names of colors vs. naming the colors of words) which is alluded to by MacLeod and Macdonald (2000) when they explained that

> The asymmetry of the interference — words interfere with color naming but colors do not interfere with word reading — suggests that reading words is more automatic (more obligatory and ballistic) than is naming colors. This makes sense in terms of our no doubt *vastly greater practice at reading*. (pp. 383–384, emphasis added)

Rather than creating vague conceptual models to explain the interference phenomenon, a simpler (i.e., more parsimonious) approach is to interpret it according to established general principles (e.g., stimulus control, generalization, reinforcement) [see Palmer, 1991]. For example, saying "red" to the word "red" and "green" to the word "green" regardless of the actual color in which the word is printed has been richly reinforced long before any experiment is carried out. When researchers then present the word "red" in the actual color of green and instruct participants to name the color of the word, the result is not surprising: a weaker response as measured, for example, by longer response latencies. Due to the much richer reinforcement history for reading words than for the odd and unusual behavior of naming the color of words (contrived by experimental psychologists), the word exerts greater stimulus control than the color of the word. This interpretation is preferable to one that appeals to mentalistic concepts such as *attention* and *long-term memory* because it rests only upon principles (e.g., reinforcement, stimulus control) that have been established independently of the phenomena to be explained (Palmer, 2003).

Moreover this interpretation could be tested by teaching very young, barely verbal, children to read color words and separately to name colors or other variations. Children could then be given the Stroop test. In this way we could actually manipulate the ultimate variables responsible for the behaviors. We could also try to construct nonhuman simulations (see e.g., Epstein, Lanza and Skinner, 1980, 1981). For example, using a matching-to-sample paradigm, pigeons (who have excellent color vision) could be taught to match color words to their corresponding colors (called symbolic matching) and, separately, to match colors to colors (called identity matching). They could be given Stroop trials at various times during training to assess the effects of color-to-color matching on the performance of words-to-color matching.

Of course, it is possible that the Stroop effect is not really that interesting from a theoretical or practical standpoint and may be more valuable as a vehicle for providing academic researchers with a fairly easy and straightforward paradigm for generating myriad research questions. Considering that hundreds of articles have been published on this phenomenon, it is telling that as recently as 1998, one of its leading researchers can say that "Even 60 years on, then, we remain rather distant from a full understanding of Stroop's (1935) deceptively simple appearing phenomenon, and the phenomenon in turn remains an important testing ground for our theories of attention (MacLeod, 1998, pp. 209). Apparently "a voluminous productivity should not be confused with progress" (Machado et al. 2000, p. 12).

Conclusion

Many experiments in psychology succeed at best in achieving only the first goal of science — description through observation and measurement (although often the variables of interest [e.g., cognitions] are not what is actually observed and measured). But because such experiments do not achieve unambiguous control of the subject matter, they fail to adequately explain the observations in terms of ultimate causal factors and they offer no solutions to practical problems. This is a far cry from the years of arduous laboratory work that typifies experimental activity in the physical sciences and in some subspecialties of psychology, such as learning and perception.

Although psychological researchers receive extensive training in research methodology, they do not receive the same degree of training or spend the countless hours in basic laboratory research that scientists in other experimental disciplines do. The result is that many psychologists may address important questions, but with a minimum of actual experimental effort or rigor. For example, in one article on compliance and the foot-in-the-door technique (Reeves, Baker, Boyd, and Cialdini, 1991) the "experiments" consisted of subjects (university students) being stopped on campus and presented with requests for a commitment of time. Those who refused were given a questionnaire. Although four "experiments" were conducted, no well-defined independent variables were systematically manipulated. In such examples, most of the researcher's time is not spent in the laboratory but, rather, with calculators and statistical tables long after the "experiment" is completed. This, in part, no doubt leads to the proliferation of research articles in psychology journals that, as Machado et al. (2000) have noted, "seems disproportional to the number of findings convincingly explained or of problems effectively solved — not to mention the number of theories that attain even a modicum of social consensus" (p. 5).

Coda

Much of psychology falls short of the purposes and goals of science. Psychologists who focus on cognitive events are interested in processes and structures that are never observed and cannot be described without resorting to metaphor. This is contrary to the hallmark of science — objective observation and measurement. Many research psychologists do not carry out thorough experimental analyses, which prevents them from fully controlling and accurately predicting the behavior of individuals. In the absence of experimental control, psychologists are not able to discover quantitative laws and principles. In the absence of such principles, psychologists are less likely to be able to offer satisfying explanations of complex phenomena. Nor are they able to offer solutions to practical behavioral problems. In fact, with the exception of the successful application of the principles of Pavlovian and operant learning to various clinical problems in programs of behavior therapy and behavior modification (applied behavior analysis), it is difficult to think of any technologies of behavior change resulting from research in nominal psychology.

Thus, with some exceptions, psychology has not fulfilled two of the purposes of science, mentioned earlier in this essay, at least not yet. Obviously, a different strategy is needed. The solution requires a serious reformulation of psychology's approach to the scientific understanding of human behavior, one that returns the focus to the relationship between behavior and its context as the primary subject matter. Those who call themselves *behavior analysts* have already made a good beginning by contributing a natural science of behavior in which basic behavioral processes have been discerned through experimental analysis. But behavior analysts have been slow to offer an empirical "account of the content domains of behavior in individual, social and cultural contexts, for instance, of perceptual and cognitive, emotional and motivated, and verbal and social behavior (Morris, 2003, p. 277), although they have offered interpretive accounts of memory (e.g., Palmer, 1991), structural regularities in verbal behavior (Palmer, 1998), cognition (Palmer, 2003), and behavioral development in children (Schlinger, 1992, 1995), to name a few.

Other programs of direct action, although not technically behavior analytic, have tackled the problem of providing a non-mediational, non-representational natural history of phenomena comprising some of the content domains in psychology. For example, Gibson's (1979) theory of direct perception treated perception as action and emphasized the mutual role of the perceiver and the environment with no spatial gaps to be filled in by mental representations. In the area of memory Watkins (1990) has offered a program of direct action, which seeks to explain remembering as a functional relationship between an experience at one point in time and behavior (remembering) at another point in time without appealing to mental representations as mediators. (For other examples of non-mediational programs of direct action in psychology and a discussion of the possible alliance between them and behavior analysis, see Morris, 2003.)

Some psychologists will bristle at the conceptual and methodological criticisms of modern psychology contained in this essay and the suggestion that psychology should (re)turn to a non-mediational, non-mechanistic, non-dualistic behavioristic approach, especially since the received view is that behaviorism is all but dead. But neither the criticisms nor the suggestion are new (see Kimble, 1996; Lee, 1988; Uttal, 2000, 2001). For instance Uttal (2000) states:

> It is also important to appreciate that behaviorism is not dead today in scientific psychology . . . . Some of us also think that it may be the main theme of the next wave of scientific psychology. (p. 45)

Such an approach, if it can provide both a natural science and a natural history of behavior, will more likely fulfill the promises of a scientific psychology. Until then, to paraphrase Robert Frost, psychology has miles to go and promises to keep.

References

Ainsworth, M., Blehar, M., Waters, E., and Wall, S. (1978). *Patterns of attachment.* Hillsdale, New Jersey: Erlbaum.

Alessi, G. (1992). Models of proximate and ultimate causation in psychology. *American Psychologist, 47*, 1359–1370.

Baars, B.J. (1986). *The cognitive revolution in psychology.* New York: Guilford.

Banich, M.T., Milham, M.P., Atchley, R.A., Cohen, N.J., Webb, A., Wszalek, T., Kramer, A.F., Liang, Z., Barad, P., Gullett, D., Shah, C., and Brown, C. (2000). Prefrontal regions play a predominant role in imposing an attentional 'set': Evidence from fMRI. *Cognitive Brain Research, 10*, 1–9.

Bernstein, D.A., Clarke–Stewart, A., Penner, L.A., Roy, E.J., and Wickens, C.D. (2000). *Psychology* (fifth edition). New York: Houghton Mifflin.

Bower, G.H. (1993). The fragmentation of psychology. *American Psychologist, 48*, 905–907.

Bushman, B.J., and Bonacci, A.M. (2002). Violence and sex impair memory for television ads. *Journal of Applied Psychology. 87*, 557–564.

Cann, A., Sherman, S.J., and Elkes, R. (1975). Effects of initial request size and timing of a second request on compliance: The foot-in-the-door and the door-in-the-face technique. *Journal of Personality and Social Psychology, 31*, 205–215.

Carlson, N.R., and Buskist, W. (1997). *Psychology* (fifth edition). Boston: Allyn and Bacon.

Chiesa, M. (1992). Radical behaviorism and scientific frameworks: From mechanistic to relational accounts. *American Psychologist, 47*, 1287–1299.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist, 49*, 997–1003.

Cohen, J.D., Dunbar, K., and McClelland, J.L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 332–361.

Epstein, R., Lanza, R.P., and Skinner, B.F. (1980). Symbolic communication between two pigeons (*Columba livia domestica*). *Science, 207*, 543–545.

Epstein, R., Lanza, R.P., and Skinner, B.F. (1981). "Self-awareness" in the pigeon. *Science, 212*, 695–696.

Fagen, J.W. (1993). Reinforcement is not enough: Learned expectations and infant behavior. *American Psychologist, 48*, 1153–1155.

Faux, S.F. (2002). Cognitive neuroscience from a behavioral perspective: A critique of chasing ghosts with Geiger counters. *The Behavior Analyst, 25*, 161–173.

Fischbach, G.D. (1992, September). Mind and brain. *Scientific American, 267*, 48–57.

Freedman, J.L., and Fraser, S.C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology, 4*, 195–202.

Gardner, H. (1985). *The mind's new science: A history of the cognitive revolution*. New York: Basic Books.

Gewirtz, J.L., and Pelaez–Nogueras, M. (1991). The attachment metaphor and the conditioning of infant separation protests. In J.L. Gewirtz and W.M. Kurtines (Eds.), *Intersections with attachment* (pp. 123–144). Hillsdale, New Jersey: Erlbaum.

Gewirtz, J.L., and Pelaez–Nogueras, M. (1992). Infant social referencing as a learned process. In S. Feinman (Ed.), *Social referencing and the social construction of reality in infancy* (pp. 151–173). New York: Plenum.

Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston: Houghton–Mifflin.

Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. New York: The Free Press.

Glenn, S.S., Ellis, J., and Greenspoon, J. (1992). On the revolutionary nature of the operant as a unit of behavioral selection. *American Psychologist, 47*, 1329–1336.

Hempel, C.G. (1966). *Philosophy of natural science*. Englewood Cliffs, New Jersey: Prentice–Hall.

Holth, P. (2001). The persistence of category mistakes in psychology. *Behavior and Philosophy, 29*, 203–219.

Huffman, K., (2002). *Psychology in action* (sixth edition). New York: Wiley.

James, L.E., and Burke, D.M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology — Learning, Memory, and Cognition, 26*, 1378–1391.

Kimble, G.A. (1996). *Psychology: The hope of a science*. Cambridge, Massachusetts: The MIT Press.

Koch, S. (1993). "Psychology," or "the psychological studies." *American Psychologist, 48*, 902–904.

Leahey. T.H. (1992). Mythical revolutions in the history of American psychology. *American Psychologist, 47*, 308–318.

Lee, V.L. (1988). *Beyond behaviorism*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Loftus, G.R. (1991). On the tyranny of hypothesis testing in the social sciences. *Contemporary Psychology, 36*, 102–105.

Loftus, G.R. (1996). Psychology will be a much better science when we change the way we analyze data. *Current Directions in Psychological Science, 5*, 161–171.

Lykken, D.T. (1991) What's wrong with Psychology, anyway? In D. Chiccetti and W. Grove (Eds.), *Thinking clearly about psychology, Volume 1* (pp. 3–39). Minneapolis: University of Minnesota Press.

Machado, A., Lourenço, O., and Silva, F.J. (2000). Facts, concepts, and theories: The shape of psychology's epistemic triangle. *Behavior and Philosophy, 28*, 1–40.

MacLeod, C.M. (1998). Training on integrated versus separated Stroop tasks: The progrssion of interference and facilitation. *Memory and Cognition, 26*, 201–211.

MacLeod, C.M., and MacDonald, P.A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Science, 4*, 383–390.

Moore, J. (1995). Some historical and conceptual relations among logical positivism, behaviorism, and cognitive psychology. In J.T. Todd and E.K. Morris (Eds.), *Modern perspectives on B.F. Skinner and contemporary behaviorism* (pp. 51–84). Westport, Connecticut: Greenwood.

Morris, E.K. (2003). Behavior analysis and a modern psychology. In K.A. Lattal and P.N. Chase (Eds.), *Behavior theory and philosophy* (pp. 275–298). New York: Kluwer Academic Press.

Palmer, D.C. (1991). A behavioral interpretation of memory. In L.J. Hayes and P.N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 261–279). Reno, Nevada: Context Press.

Palmer, D.C. (1998). The speaker as listener: The interpretation of structural regularities in verbal behavior. *The Analysis of Verbal Behavior, 15*, 3–16.

Palmer, D.C. (2000). Chomsky's nativism: A critical review. *The Analysis of Verbal Behavior, 17*, 39–50.

Palmer, D.C. (2003). Cognition. In K.A. Lattal and P.N. Chase (Eds.), *Behavior theory and philosophy* (pp. 167–185). New York: Kluwer Academic Press.

Palmer, D.C., and Donahoe, J.W. (1992). Essentialism and selectionism in cognitive science and behavior analysis. *American Psychologist, 47*, 1344–1358.

Pinker, S. (1994). *The language instinct*. New York: HarperCollins.

Reeves, R.A., Baker, G.A., Boyd, J.G., and Cialdini, R.B. (1991). The door-in-the-face technique: Reciprocal concessions vs. self-presentational explanations. *Journal of Social Behavior and Personality*, 6, 545–558.

Russell, B. (1961). *History of Western philosophy* (second edition). London: Allen and Unwin Ltd.

Schlinger, H.D. (1992). Theory in behavior analysis: An application to child development. *American Psychologist, 47*, 1396–1410.

Schlinger, H.D. (1993). Learned expectancies are not adequate scientific explanations. *American Psychologist, 48*, 1155–1156.

Schlinger, H.D. (1995). *A behavior-analytic view of child development*. New York: Plenum.

Schlinger, H.D. (1998). Of planets and cognitions: The use of deductive inference in the natural sciences and psychology. *The Skeptical Inquirer, 22*, 49–51.

Schwartz, B. (1989). *Psychology of learning and behavior*. New York: Norton.

Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson and M.D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213–245). Hillsdale, New Jersey: Erlbaum.

Skinner, B.F. (1956). A case history in scientific method. *American Psychologist, 11*, 221–233.

Skinner, B.F. (1974). *About behaviorism*. New York: Knopf.

Slife, B.D., and Williams, R.N. (1997). Toward a theoretical psychology: Should a subdiscipline be formally recognized? *American Psychologist, 52*, 117–129.

Solso, R.L. (2001). *Cognitive psychology* (sixth edition). Boston: Allyn and Bacon.

Staats, A.W. (1983). *Psychology's crisis of disunity: Philosophy and method for a unified science*. New York: Praeger.

Staats, A.W. (1991). Unified positivism and unification psychology: Fad or new field? *American Psychologist, 46*, 899–912.

Staats, A.W. (1999). Unifying psychology requires new infrastructure, theory, method, and a research agenda. *Review of General Psychology, 3*, 3–13.

Stanovich, K.E. (1998). *How to think straight about psychology* (fifth edition). New York: Longman.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Uttal, W.R. (2000). *The war between mentalism and behaviorism: On the accessibility of mental processes*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Uttal, W.R. (2001). A credo for a revitalized behaviorism: Characteristics and emerging principles. *Behavioural Processes, 54*, 5–10.

Viney, W. (1996). Disunity in psychology and other sciences: The network or the block universe. *Journal of Mind and Behavior, 17*, 31–44.

Vyse, S.A. (1997). *Believing in magic: The psychology of superstition*. New York: Oxford University Press.

Watkins, M.J. (1990). Mediationism and the obfuscation of memory. *American Psychologist, 45*, 328–335.

Watson, J.B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*, 158–177.

Wilkinson, L. and Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594–604.

Zeiler, M.D. (1986). Behavioral units: A historical introduction. In T. Thompson and M.D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 1–12). Hillsdale, New Jersey: Erlbaum.