

Elaborate reflections on a simple manifesto

RICHARD M. MCFALL

Indiana University, Bloomington

Abstract

In an earlier article (McFall, 1991), I urged clinical psychologists to work toward the goal of integrating science and practice and proposed the adoption of a principle and two corollaries aimed at achieving this goal. In general, I argued that all aspects of clinical psychology must be guided by the highest scientific and ethical standards, that clinical practice be limited to empirically supported procedures, and that clinical training be devoted to producing clinical scientists. In the present article, I elaborate and defend these points by offering reflections on a number of submitted questions provoked by the earlier article. I address four major issues: the philosophical foundations for a scientific epistemology, the implications of this epistemology for clinical practice, the implications for clinical training, and the likely impact of adopting this epistemology on the field of clinical psychology.

Key words: Scientific epistemology, Standards, Ethics, Therapy, Training

The ideas which are here expressed so laboriously are extremely simple and should be obvious. The difficulty lies, not in the new ideas, but in escaping from the old ones, which ramify, for those brought up as most of us have been, into every corner of our minds.

—John Maynard Keynes

Background

This was supposed to be a straightforward article. The editor, David Smith, would invite members of the American Association of Applied and Preventive Psychology (AAAPP) and the Society for a Science of Clinical Psychology (SSCP) to submit questions concerning my earlier article, “Manifesto for a Science of Clinical Psychology” (McFall, 1991). I then would respond in a Q & A format. But when Smith eventually sent me 14 single-spaced pages of questions, I realized that the original plan was in trouble on three counts. First, many submissions really were assertions and counterarguments thinly disguised as questions. Second, even if my responses consumed little more space than the original questions, the Q & A format would force me beyond my page limit. Third, if I responded to the far-ranging queries individually, the article would be so scattered and tedious that few readers would find it interesting or illuminating.

So I took a different tack. First, I sorted the questions into logical categories based on content similarity, thereby distilling the 14 pages of questions into four major areas of concern.

Then, for each submission I highlighted the specific point and, if needed, edited it to make it succinct. Finally, I tried to address each of the four major issues in a logical, integrated, and constructive way. I wanted not only to provide a conceptual framework that clarified the confusion and reduced the consternation occasioned by the Manifesto, but also to go beyond what had been said before.¹

This is my second reprise of the Manifesto. In 1996, I responded to a critique by Donald R. Peterson. Our articles were published together in *Applied & Preventive Psychology*. I took that occasion to extend the original principles, using his criticisms as pointers to places where I had been cryptic or ambiguous. (A summary of the original principles and their subsequent amplification is appended here.) Judging from the questions submitted for the present article, my 1996 effort had little impact on some readers’ misgivings about the original Manifesto.

No one has been more surprised than I by the reactions to the Manifesto since its publication in *The Clinical Psychologist*. I initially wrote it as my presidential address to SSCP, in the summer of 1990, while holed up alone in a cabin on a lake in Michigan. With no one to give me feedback, I was unsure how it would be received. On the one hand, I worried that my points were so simple and self-evident that I might be restat-

¹ Throughout the present article, as in the original Manifesto, I focus on clinical psychology. Many of the issues apply to other areas and specialties as well. Counseling psychologists simply should substitute the word “counseling” for “clinical” throughout. The translation to other areas of human psychology—biological, cognitive, developmental, industrial/organizational, social—may take slightly more effort. In general, the article should provide food for thought to any professional who intervenes in other people’s business with the goal of improving things (e.g., business consultants, clergy, lawyers, physicians, politicians, social workers, teachers, etc.).

Send correspondence and reprint requests to: Richard M. McFall, Department of Psychology, Indiana University, Bloomington, IN 47405. E-mail: mcfall@indiana.edu

ing the obvious. The ideas certainly weren't original. I had learned them as a graduate student in the clinical program at The Ohio State University from my mentors, Julian B. Rotter and George A. Kelly. Indeed, Rotter (1971) had published similar ideas in *The Clinical Psychologist* two decades earlier. On the other hand, when I observed the current state of affairs in clinical psychology, it seemed that the majority of clinical psychologists either were unfamiliar with these important principles or were ignoring them. In either case, the time seemed ripe for SSCP—an organization dedicated to the advancement of clinical psychology as a science—to take a more active role in articulating and promoting such principles. Even if I were preaching to the choir, the Manifesto might help ensure that everyone was singing from the same hymnal.

Someone has asked if I have had any afterthoughts about the Manifesto or if I would write it differently were I to do it today. To me, the Manifesto's principles seem as simple, self-evident, and valid today as they did when I wrote them a decade ago or even when I first encountered them as a graduate student almost 40 years ago. It is the nature of principles that they do not change; that is what makes them principles! But I could have stated them more clearly or explained them more cogently. I am pleased to have this opportunity to try again.

The arguments in the original Manifesto were stated as absolute moral imperatives, with frequent use of the words "ought," "should," and "must." In retrospect, this all-or-none language was a source of some confusion. It was my intent in the Manifesto to distinguish between the present standards of practice in psychology and the standards of practice that would prevail if our actions were guided by the highest scientific and ethical principles. The choice between these two sets of standards seemed so clear to me that I neglected to walk through the underlying epistemological rationale for this preference or to discuss in detail the practical implications of this choice for clinical practice. In the present article, I devote considerable space to correcting this oversight by explaining in detail how my absolutist stance regarding the adoption of a scientific epistemology in clinical psychology is logically consistent with a less absolute, more probabilistic approach toward the development, evaluation, and application of clinical theories and methods.

My earlier neglect was reflected in many of the questions I received for the present article. The two most common themes were (a) the philosophical justification for the Cardinal Principle (i.e., "scientific clinical psychology is the only legitimate and acceptable form of clinical psychology") and (b) the real-world implications of adopting this principle and its First Corollary (i.e., that psychologists should offer only those services that have been shown empirically to be safe and effective). The other two major themes were (c) the implications for clinical training (the Manifesto's Second Corollary) and (d) the possible impact on the field of adopting such principles. In the following section, I tackle these four themes sequentially, first presenting the essence of all the questions on each theme and then offering an integrated response.

Q & A DIALOGUE: WHAT THE MANIFESTO DOES AND DOES NOT SAY

Issue 1: Questions on Philosophy of Science

- How scientific is scientific? If the support has to be massive and airtight, hands will be tied and little will get done.
- At what level should one be scientific? What is the precise operational definition of "validated scientifically"?
- What are the demarcation criteria that enable us to tell the difference between scientific clinical psychology and other kinds, and are there any proxies (shortcuts) by which the public can distinguish?
- Your thesis is based on a logical positivist position, which has been criticized roundly; what about other philosophies of science, which would set different standards of quality?
- By what methods, designs, and standards do you establish the scientific status of clinical practices?
- Science is limited to efficient causal accounts (i.e., explanations in terms of antecedent events). What about final causality (i.e., explanations based on purpose, teleological matters, and free will)?
- Isn't science just a way of knowing and learning? If so, shouldn't clinicians' effective behaviors be reinforced and ineffective behaviors be punished, leading to effective psychotherapy in the absence of formal scientific input?

Reflections on Issue 1: Adopting a Scientific Epistemology

The Manifesto insisted that clinical psychology adopt a scientific epistemology. To address questions on this issue, I need to explain what I mean by a "scientific epistemology." "Epistemology" refers to the origins, nature, methods, and limits of knowledge, to how we "know" anything or decide what is "true." Historians of science (e.g., Boorstin, 1983; Popper, 1962) have described numerous ways that humans throughout the ages have decided what is "true."² Science is only one of many epistemological approaches. Over the past 300 years, a scientific approach has been extremely fruitful in, say, astronomy, biology, chemistry, geology, medicine, physics, and zoology. Many of us believe that a rigorous scientific epistemology would be equally fruitful in clinical psychology.

²For example, victors in battle are judged to be on the side of "truth" ("might makes right"); "truth" is decreed by authorities or experts, such as kings, priests, shamans, or scholars; certain individuals, such as Mohammad, are selected by God as conduits for the "revealed truth"; knowledge is inborn, and so "truth" can be discovered through deductive logic or intuition; "truth" can be induced from observation or experience; "truth" is revealed to those who earn it through pain, privation, or good works; "truth" is revealed to the pure of heart, the childlike, the devout, or the spiritual through divine revelation, inspiration, or meditation; "truth" is evident to juries of neutral nonexperts who listen to argument and weigh the evidence; etc.

Theoretical ideas about “truth” are not the exclusive domain of science. Such ideas are the products of creative thinking and can come from anyone or anywhere. But a scientific epistemology provides a powerful array of conceptual, empirical, and quantitative tools for investigating the relative merit of these competing ideas and for pursuing their logical and practical implications. A scientific epistemology is not an “automatic inference machine” (Meehl, 1971); it is not a single, rote, “paint-by-the-numbers” method of determining “truth.” Rather, it is a general perspective with a related set of methods, which, if applied consistently and conscientiously, increases our chances of homing in on “truth.” Indeed, given the inherent difficulties of studying complex psychological phenomena (Meehl, 1978), a scientific epistemology seems to offer the only hope for achieving solid advances in clinical psychology.

Scientists build theoretical models of reality and then test these models (McFall & Townsend, 1998). Scientific models are simplified representations of reality, approximations to “ultimate truth.” Today’s “truth” always is at risk of being revised or replaced as a result of tomorrow’s new evidence. Thus, “knowledge” and “truth” are abstract constructions, not absolute realities. It is an error to reify our models. (To emphasize this point, I’ve been putting words such as “truth” in quotation marks—an annoying practice I’ll stop now.)

Scientists are not the only ones who build theoretical models of reality. So, how do scientific models differ from pseudoscientific or nonscientific models? One of the most critical differences is that scientific theories generate risky predictions that are amenable to objective, replicable tests. A “risky” prediction not only is capable of being disconfirmed or falsified, but also goes beyond the observations that gave rise to the theory in the first place (Feynman, 1985; Popper, 1962). A theory that provides a new and better “fit” to nature, relative to competing models, reduces uncertainty (Shannon & Weaver, 1949); it increases the accuracy of predictions over what could have been achieved without it; it has “incremental validity” (Meehl, 1959; Sechrest, 1963).

Ordinarily, a theory’s validity, or scientific status, is not judged in an all-or-none manner. As evidence is gathered over time, the theory’s status is enhanced or diminished by degrees. Initially, we may say that a theory was “supported” by the evidence (or was not disconfirmed); with replication and further testing, we may say that it was “validated”; eventually, after it has withstood repeated tests from multiple angles, we may act as though the theory were absolutely true, even though it still is just a model of truth.

Contemporary scientific models tend to be relative and probabilistic rather than absolute and deterministic (Gigerenzer & Murray, 1987; Gigerenzer et al., 1989). The so-called hard sciences have made this transition; psychological science needs to do so as well. Since Heisenberg, for example, physicists have accepted the impossibility of measuring simultaneously the exact position and momentum of an atom. Given such indeterminacies, physicists now focus on assess-

ing the “probability distributions” for selected characteristics of many atoms—rather than single atoms—observed under specific conditions.

This shift from deterministic to probabilistic models reflects a realization that we live in a universe in which specific events are affected by random or chance inputs. The variability of observed outcomes, as depicted in probability distributions, no longer can be dismissed as an artifact of incomplete theories and imperfect measures; rather, such variability now must be accepted as an inherent, irreducible characteristic of nature. This means that nature is represented inadequately by a single value; it is represented more faithfully by an array of outcomes, each associated with a probability or likelihood. The goal of our theories, then, is to model these probability distributions and predict how they will change under varying conditions. Such contextual dependencies are expressed as conditional probability distributions.

This view of contemporary scientific epistemology, with its emphasis on probabilistic models, has important ramifications for the philosophical foundations of theory and practice in clinical psychology. First, it explains why the goal of predicting specific events with precision is illusory. Second, it explains why idiographic approaches to clinical assessment and intervention in psychology are unrealistic. In an individual case, any predictions regarding, say, the implications of psychological test results or the effectiveness of a treatment method are stated most realistically in probabilistic terms. Such probabilistic statements inevitably are nomothetic. Like the meteorologist forecasting tomorrow’s weather (“There is a 60% chance of rain.”), the best a clinician can hope to achieve in a given case is to say, for example, “based on these psychological assessments, I estimate that x has y probability of occurring” or “for a client with this problem, this intervention has y probability of achieving x results.” These are nomothetic statements, which could range from wild guesses, based on pure intuition, to empirically grounded likelihood estimates, based on the observed probability distributions of prior outcomes.

The accuracy of probabilistic statements is affected by at least two major factors: (a) the degree of similarity between the predicted cases and the sample of prior cases on which the predictions were based, and (b) the size of the prior sample. Thus, predictive accuracy can be improved by increasing the size and representativeness of the samples on which we base our conditional probability estimates. Most psychologists understand what a “large sample” is. But what is a “representative sample” (Maher, 1978)? The more similar the predictive sample is to the target sample, the more representative it is. So how do we determine the “similarity”?

First, because the critical dimensions of similarity are not self-evident, we must choose which features we’ll attend to and which we’ll ignore. Every choice represents an implicit or explicit theoretical conjecture about the underlying structure of nature (Popper, 1962). For instance, a whale and a human are quite dissimilar on the surface but are regarded

within zoological theory as more similar than, say, a whale and a shark because they share critical features that define the theoretical category of mammal. In an analogous way, clinical psychologists must decide, on theoretical grounds, which abstracted features, or dimensions, are most important for judging the degree of similarity among clinical phenomena. Every estimate of conditional probabilities implies an underlying theoretical model of critical similarities and differences. Theoretical models underlie even the simplest of clinical generalizations.

Although all predictions necessarily are probabilistic and theoretical, the actual outcomes in individual instances are definitive and concrete. When estimating the likelihood of two mutually exclusive outcomes, for instance, we would say that outcome *a* will occur *x* percent of the time, and that outcome *b* will occur $100 - x$ percent of the time. But in any given instance, either *a* or *b* actually occurs. Thus, the meteorologist may estimate that the probability of rain tomorrow is 60%, but in fact it either rains or it doesn't.³ Similarly, your odds of winning a lottery may be one in ten million, but your ticket either wins or loses. Probabilistic predictions of continuously scaled variables (e.g., the meteorologist predicts a high of 72°F tomorrow) are only slightly more complicated. We transform these variables into an array of "bins" of possible outcomes, *a* through *n* and estimate the probability associated with each bin, with the sum of all probabilities = 1. Once again, however, the outcome actually observed on each occasion falls into only one bin.

So if predictions are probabilistic but specific outcomes are definitive, how can we assess the accuracy of our predictions? And, by implication, how do we test the validity of the theories on which these predictions are based? Obviously, we cannot possibly do this based on the outcome of a single case! An adequate test requires a large sample of paired predictions and outcomes; this allows us to assess quantitatively the "fit" between the observed and predicted distributions of outcomes. The larger the sample of paired predictions and observations, the more confidence we can have in our conclusions about the predictive accuracy of our theories and methods. These fundamental concepts of probability theory, sampling, and the law of large numbers are familiar to anyone who has taken a basic statistics course. Nevertheless, they often seem to be ignored in clinical practice.

We have reviewed some of the quantitative reasons why adopting a scientific epistemology requires that psychologists take a probabilistic, nomothetic approach to clinical theory and practice. There are logical reasons for doing so as well (McFall & McDonel, 1986). As we've noted, all rational predictions imply some theoretical conception of the best way to classify events (e.g., clients, problems, interventions,

outcomes, etc). It is impossible to create such classification systems from scratch with each new client. Thus, it is logically impossible to "treat each client as unique." To process unique events, we must view them through the lens of our existing conceptual structure.

Ironically, some clinicians who claim to approach clients idiographically also claim that they rely on their personal clinical experience to decide what each client needs and how best to serve those needs (Peterson, 1996). "Personal clinical experience" is another name for "generalizations based on prior cases." Thus, relying on "clinical experience" is a nomothetic approach in disguise. The problem with knowledge claims based on "personal clinical experience" is that these represent subjective or "intuitive" knowledge that typically is not stated explicitly, and so its validity cannot be tested. This is why such claims are viewed skeptically within a scientific epistemology. To be consistent with the Cardinal Principle, clinicians must state their claims of experiential insights in a form that leads to risky and testable predictions and must be willing to accept the verdict of the empirical evidence regarding the validity of their claims.

The preceding explication of "a scientific epistemology" sheds light on many of the specific questions I received on the first issue (i.e., philosophy of science questions). For example, the discussion of quantitative and qualitative standards for assessing the validity of clinical theories and methods addressed questions such as "How scientific is scientific?" "What are the criteria for 'scientific'?" "What is the precise operational definition of 'scientifically validated'?" Specifically, validity judgments seldom are all-or-none but are incremental. A theory or method is judged to be valid to the degree that it reliably reduces the uncertainty, or error, of our predictions across large samples of paired predictions and observations. Essentially, this is a Bayesian conception: incremental validity is defined quantitatively as the magnitude of the difference in accuracy between our a priori (pretheory) and a posteriori (posttheory) predictions (cf. McFall & Townsend, 1998; McFall & Treat, 1999).

The Manifesto's Cardinal Principle, with its call for a scientific epistemology, was a postulate, an assumption, a "given." It was the foundation for the logical argument that followed. Postulates are not debatable; one either accepts them (at least provisionally) or not. If clinical psychologists cannot agree on their fundamental postulates, they should not attempt to obscure such differences. It only leads to confusion, for example, if the same label—"clinical psychology"—is applied to two or more incompatible epistemological systems.

To avoid such confusion, the Cardinal Principle unequivocally identified *which* type of clinical psychology was being discussed and differentiated it from other types of clinical psychology that rely on other ways of knowing. It called for a total commitment to a scientific epistemology. This absolutist position undoubtedly made some readers uneasy, but clinical psychologists have suffered too long from "scientism" (i.e., longing to be considered scientists while failing

³Of course, things may not be quite this simple in reality. It might rain in my back yard but not my neighbor's; there might be mist but not raindrops. Meteorologists handle such ambiguities by recording measurable precipitation in a rain gauge at an "official" weather station.

to act consistently as scientists). For example, there are too many “two-headed clinicians” (i.e., psychologists who may think critically in their laboratory but uncritically in the therapy room). If we are serious about advancing knowledge regarding the etiology, course, concomitants, measurement, prevention, and amelioration of psychological distress and disorder, then our only plausible option is to adopt a scientific epistemology. Other disciplines that have made an absolute commitment to a scientific epistemology have achieved significant advances in knowledge. Similar advances in clinical psychology await a similar commitment.

The strength of a scientific epistemology is that it circumvents most of the pitfalls of other systems for resolving questions of truth. For example, one of the submitted queries for this article asked whether natural contingencies might not shape the behavior of clinical practitioners, producing effective methods of psychotherapy in the absence of formal scientific input. This question must be answered by a question: How could we possibly know, with any confidence, whether the therapies produced by natural contingencies actually are effective or whether therapists simply believe them to be so? The way to resolve such a question is to conduct controlled, empirical tests of the therapies.

Another questioner asked whether other philosophies of science might not lead to other standards of practice. This, too, invites a question in return: What alternative philosophy of science does the questioner have in mind? How does this alternative’s record for resolving questions of truth and for advancing knowledge compare with that of the scientific approach I am advocating?

Another questioner expressed concern that a scientific epistemology excludes consideration of “final causes” and “free will.” Granted, a scientific approach requires, at a minimum, that concepts be defined with sufficient precision that judges can apply them reliably (i.e., “know them when they see them”). If “final cause” and “free will” satisfy this requirement, they can be studied scientifically; if not, I see no satisfactory way to evaluate their validity or utility.

This raises one final point. Clinical psychologists cannot decide to accept or reject a scientific epistemology simply because it does or does not yield results that confirm their pet theory or outlook on life. Indeed, the strength of a scientific epistemology is that it helps us overcome our prejudices and blind spots. Of all truth-telling strategies, a scientific approach is the one that, in principle, is least susceptible to error due to the corrupting influences of authority, belief, intuition, bias, preference, etc.

This concludes my reflections on Issue 1. The remainder of the article assumes that, for the sake of argument, readers are willing to accept, at least provisionally, the Cardinal Principle as elaborated here. Issue 2 concerns the Cardinal Principle’s implications for clinical practice. The most immediate implication (covered in the Manifesto’s First Corollary) is that psychologists should offer only those services that clearly are safe and effective. Judging from the following

questions, such “real world” implications require further elaboration and justification.

Issue 2: Questions on Implications for Practice

- By what criteria of “legitimacy” and “acceptability” did you derive the First Corollary?
- How do you justify generalizing scientific research findings to applied practice when the data supporting doing so are still quite sparse?
- How do you deal with the complaint that research findings are based on narrow patient samples, unrepresentative therapists, etc.?
- How exact is “exact”? Given the variable degree of empirical support for many services that are in demand, this is not mere hairsplitting.
- How should one proceed in the absence of data?
- If a student asked you whether s/he should take notes during or after a therapy session, would you suggest looking at the literature, doing a study, tossing a coin, or going with instinct?
- How do we make routine decisions on matters not addressed in the literature (e.g., choice of furniture, wall decorations, time of day for appointments, etc.)?
- How do we deal with patients whose problems are too complex to have any treatment available that would conform to acceptable scientific standards?
- Exactly which techniques, applied to which problems, experienced by which clients have been scientifically validated at this time, and what proportion of the services currently being offered fall into the above categories?
- Rather than classifying techniques as “supported” or not, wouldn’t it be better to specify the level of confidence that can be placed in the use of a given technique for a specified client group with an array of problems as identified by subsets of clinicians?
- How do you deal with the demand for services where there are no established remedies?
- Isn’t the hallmark of a profession, as opposed to a trade, that its members are capable of proceeding effectively in the absence of specific algorithmic protocols? Don’t we expect professionals to use good judgment and to manage uncertainty better than nonprofessionals do?
- Medical liability is mitigated when it can be shown that the implicated care followed “routine and accepted practice.” What is wrong with allowing the same thing in applied psychology, in addition to “empirically supported practice”?
- Should we reserve the right to practice some things of unproven validity until the science catches up?
- Rather than refusing to offer treatment, in the absence of evidence, why not give the treatment and turn it into a study?

- Would treating every case as an $N = 1$ study qualify as conforming one's practice to the Manifesto?
- Can a clinician be established as a proven treatment?
- Isn't clinical practice self-corrective to some degree, forcing practitioners to abandon ineffective treatments because these undermine their reputations and ruin their livelihoods?
- Does "artistry" play any role?
- Must studies that attribute considerable benefit to placebo and common nonspecific factors be discarded?
- Doesn't complete informed consent undermine important placebo contributions to efficacious therapies?
- I applaud the Cardinal Principle, but the First Corollary is too impractical and idealistic to be taken seriously by most practitioners. Could it (should it) be toned down?

Reflections on Issue 2: Implications of a Probabilistic Scientific Epistemology for Decision-Making

Let's begin by dispensing with a common misconception regarding the relationship between science and practice. For some reason, many clinical psychologists think of basic science and clinical practice as polar opposites on a single dimension. This is an inaccurate and overly simplistic view; it should be replaced by a two-dimensional representation. In the new model, the vertical axis represents the epistemological dimension, with "science" anchoring the top and "non-science" anchoring the bottom. The horizontal axis represents the activity-context dimension, with "basic" anchoring the left end and "applied" anchoring the right end. In this expanded view, science and practice no longer are contrasts; conceptually, they are orthogonal. Although they seem to be correlated currently, with an overrepresentation of psychologists in the basic science and applied nonscience quadrants, this need not be the case.

According to this two-dimensional model, all psychologists in the upper half would behave as scientists, regardless of the particular activity context in which they were functioning at the time, from the most basic research lab to the most applied treatment program. These psychologists might function in different contexts at different times but always would employ a scientific epistemology. In contrast, psychologists in the lower half might function in the same range of activity contexts but with an unscientific epistemology. The Cardinal Principle asserts that clinical psychology, as a discipline, should be limited to the upper half of this two-dimensional model. The following discussion focuses on clinical psychology in the upper right ("applied science") quadrant of this model.

Any psychologist who offers to perform clinical services for a client (outside of a controlled experimental context, which we will discuss in a moment) implicitly or explicitly is making both a claim and a prediction. First, the psychologist is claiming to have special knowledge and competence in the assessment and treatment of the client's psychological

difficulties. This claim is reinforced in many subtle ways, including the use of the title "Doctor"; wall displays of framed degrees, state licenses, and professional affiliations; yellow-page ads and professional offices; and the collection of fees for the services. The public has been conditioned to assume that all licensed professionals—from barbers to brain surgeons, from plumbers to pilots—know more than nonprofessionals about their particular area of expertise, and that they can apply their specialized knowledge skillfully, thereby providing services of higher quality than might be expected from nonprofessionals. Second, the psychologist also is making an implicit or explicit probabilistic prediction that, on average, clients who avail themselves of the professional services will benefit from having done so. The psychologist is saying, in essence, "I expect that my services may be able to help you."

Meanwhile, clients suffering from psychological problems are inclined to grasp at straws; persons in desperate straits tend to be susceptible to illusions. Clients who contract for psychological services obviously believe—or at least hope—that there is a reasonable chance that their lives will improve as a consequence. Thus, clinicians who offer psychological services not only imply benefits but also, if they fail to discuss likely results, leave clients' illusions about outcomes intact. Most clinicians know better than to predict a specific outcome in an individual case, but few provide the kind of valid probabilistic information that clients need in order to have reasonable expectations and to make rational decisions.

Any reader tempted to dispute the view that psychologists who offer services are making implicit claims and predictions should consider the alternative. Would clinicians be willing to stipulate in writing that they do not have any special knowledge, and that their services are not likely to be beneficial? Any clinician willing to agree to such stipulations would be out of business in short order. Most clinicians would not agree to these stipulations because they honestly believe that they do have special knowledge and skill, and that their services are beneficial on the whole. So, to offer services is to make implicit claims and predictions.

But to be consistent with the Cardinal Principle (i.e., that clinical psychologists should function solely within the upper half of the two-dimensional model), these implicit claims and predictions must be viewed with scientific skepticism. Are they valid? How do we know? Thus, the First Corollary declares that psychological services should not be administered to the public (except under strict experimental control) until they have satisfied these four minimal criteria:

1. The exact nature of the service must be described clearly.
2. The claimed benefits of the services must be stated explicitly.
3. These claimed benefits must be validated scientifically.

4. Possible negative side effects that might outweigh any benefits must be ruled out empirically.

The underlying logic should be obvious. First, unless clinicians provide a clear description of the services they intend to provide—and why—in each case, the validity of their methods and their underlying theories cannot possibly be evaluated. A vague and inexact description leaves too much wiggle room. Second, unless clinicians provide a clear probabilistic prediction regarding the expected effects of their services, it is impossible to assess the incremental validity and efficacy of the services for particular clients and problems. In short, clinicians who refuse to say what they are doing and why and who refuse to say what can be expected from their services simply cannot be held accountable. Third, a clinician who offers services (thereby making implicit claims) cannot possibly be functioning as an applied scientist unless these services (and claims) have been tested in ways that satisfy scientific standards for incremental validity, as discussed under Issue 1. Fourth, scientific evaluations of the effects of psychological services must take into consideration both the positive and negative outcomes, expected and unexpected. To arrive at a summary judgment, these multiple outcomes should be assigned weights corresponding to their valence and relative importance to the clients.

One questioner (not included above) focused specifically on the fourth criterion: “Why restrict it to possible negative side effects that only ‘outweigh’ positive effects? Any negative side effects ought to be of concern, even those that simply subtract from, but don’t completely offset, the positive effects. This issue is concerned with the ‘net’ benefit.” This questioner helps sharpen the issue. Indeed, we must consider the cumulative weight of all positive effects minus the cumulative weight of all negative effects. As in any cost-benefit analysis, the net difference is what concerns us. Sometimes the sum of many small negative effects can outweigh the sum of a few major benefits. One way to increase the net benefit of an intervention, for example, is to reduce the likelihood of negative side effects.

The First Corollary provides for one exception: untested services may be administered as long as they are identified in advance as experimental, they are administered under controlled conditions permitting reasonable scientific inferences, and clients are informed fully and consent to be participants in an experimental test of the services. Informed consent includes the knowledge that, because the procedures are experimental, no claims or empirically based predictions of outcome can be made, and that some clients will be assigned to a “control” treatment.

The First Corollary rests on two pillars—one logical, one ethical. The logical pillar, deduced from the Cardinal Principle, asserts that, if we are to act scientifically, we must do all we can to ensure that our decisions and actions in clinical practice are guided by and consistent with the best scientific theories and empirical evidence. The ethical pillar asserts

that practitioners’ first obligation is to do no harm. They knowingly must not cause physical or emotional pain and suffering, must not make false or misleading claims or promises, must obtain full informed consent from clients before taking actions that will affect them, and must not engage in other dishonest or self-serving acts that might harm the client. This familiar injunction is codified in ethical guidelines published by the American Psychological Association (1992) and the National Academy of Sciences (1995), in legal codes (e.g., Indiana Code 25–33), and in many religious texts.

Judging from the questions I received, many psychologists are willing to endorse these logical and ethical injunctions in the abstract but are genuinely concerned that, if these “ideal” injunctions were enforced in “real life,” they would force clinicians to abandon practice altogether. As one questioner put it, “The First Corollary is too impractical and idealistic to be taken seriously by most practitioners. Could it (should it) be toned down?”

This concern must be faced head-on. Clinicians who offer services for which there is no good scientific evidence of validity are violating both the logical and ethical tenets of scientific clinical psychology. They are “winging it,” which is unacceptable. They may argue that their services are backed by “clinical experience,” that their treatments are consistent with “routine and accepted practices,” or that, at a minimum, their services yield placebo effects or offer clients “hope.” But none of these justifications for “business as usual” satisfies scientific standards.

First, practitioners who “wing it” are misleading their clients by pretending to know more than they do; this violates accepted ethical and legal standards for truth in advertising. Second, these practitioners may be putting their clients at risk of unknown side effects—effects less severe, perhaps, than those associated with medical interventions, but effects, nonetheless, that clients might find unacceptable if advised of them in advance. One obvious example is the hidden opportunity costs in time, money, and emotional resources of participating in psychotherapy. Might clients spend these same resources on other options yielding greater benefits? Third, practitioners who “wing it” are experimenting (fooling around, meddling) without first obtaining clients’ full and informed consent; this is unethical. Fourth, practitioners who offer intuitive and unsystematic interventions have no sound basis for drawing inferences or estimating outcomes. In short, these practitioners are engaging in precisely the kind of unscientific, unethical clinical practice that the Manifesto challenges. This challenge cannot be “toned down” without undermining the essential tenets of scientific clinical psychology.

Some questioners argued that requiring informed consent may undermine the potential therapeutic benefits of placebo effects. Setting aside for the moment any ethical concerns this may raise about the use of deceit or the disregard of clients’ rights, let’s look more closely at the underlying logic

of this appeal to preserving “placebo effects” in psychotherapy. In the first place, it presumes that psychological interventions actually yield beneficial placebo effects. But how can this be so? The term “placebo” has no specific referent in psychology; we have no true placebos; there is no equivalent to sugar pills or inert interventions. Thus, the term is being used metaphorically in this argument to represent all of the vague, unspecified, residual aspects of psychotherapy that may explain why some clients improve more than we would have expected on the basis of the “official” treatment alone. Almost by definition, these “unofficial” therapeutic ingredients must remain unspecified and untested. If we could identify them, assess their effects experimentally in clinical trials, and demonstrate empirically that they had incremental validity, then we undoubtedly would incorporate them in our treatment protocols. Once they were explicit treatment components, they no longer would qualify as “placebos.”

If, indeed, there is something beneficial about being a client in psychotherapy, something unspecified that we do not understand, then it behooves us to investigate this puzzle scientifically rather than shrouding it in a mysterious cloak of “placebo effects.” Furthermore, if it can be shown that unspecified aspects of psychotherapy actually yield genuine psychological benefits, then I see no reason why we should not share this hopeful fact with clients on our informed consent forms. We might say, “Many individuals seem to derive benefits from the simple act of deciding to participate in psychotherapy. Research has shown, for instance, that this decision alone leads to x amount of improvement in y percent of the individuals with z problem.” I am not aware of any evidence that such an honest revelation diminishes “placebo effects”; if anything, telling clients about the research evidence should enhance such effects. Of course, if there is no clear research evidence of such benefits, the “placebo effects” argument against requiring informed consent is a presumptuous and hollow argument.

To the extent that psychologists use the term “placebo effects” metaphorically to refer to the unknown causes of empirically demonstrated positive changes associated with a particular form of psychotherapy, I have no serious objection other than that it may be misleading. Used in this way, it serves as a “place holder” for yet-to-be-determined active ingredients. It is reasonable for clinicians to employ such empirically supported interventions, even though the mechanisms responsible for the effects remain a mystery. In medicine, for example, aspirin was accepted as a valid headache remedy long before anyone knew how or why it worked. In psychology, too, we should be open-minded about any intervention that works, even though we don’t know how or why. Rotter (1971) even suggested that a therapist might be viewed as an empirically supported treatment, provided that the therapist consistently yielded positive effects, even though the therapist might not be able to explain how s/he achieved these effects. From this perspective, “placebo effects” and “therapist artistry” need not fall outside the realm of science.

But as scientists, we should object when we see psychologists appealing to “placebo effects” and “therapist artistry” to justify their continued use of methods that lack empirical support. These appeals seem to take the following illogical form. Research has shown that some psychological interventions/therapists yield positive effects, although the scientific explanation for these effects remains a mystery. As a therapist, I administer psychological interventions that lack a clear explanation. Therefore, I also am likely to produce mysterious, unspecified positive effects. This faulty syllogism is not a satisfactory substitute for scientific evidence.

Finally, psychologists’ appeal to “placebo effects” and “therapist artistry” trivializes what clinical psychology now has to offer. Thanks to the persistent efforts of psychological scientists over the years, we no longer need to base our professional activities on such weak foundations. Empirically supported treatments now are available for the vast majority of the psychological disorders encountered most frequently by practitioners. Thus, the First Corollary’s injunction against offering unsupported services does not mean that all applied clinical activity would have to be abandoned. On the contrary, practitioners now can rely on empirical research to reduce their uncertainty about which clinical methods have incremental validity for particular clinical disorders. A detailed summary of this extensive research literature (as requested by one questioner) is beyond the scope of the present article. Fortunately, such information is becoming more readily available (e.g., Chambless, 1995; Chambless et al., 1996; Practice Guidelines Coalition, 1999). Given the rapidly evolving state of the evidence, however, there is no substitute for staying abreast of the literature oneself. For this reason, one of the hallmarks of an applied scientist is the ability to read, evaluate, interpret, and be guided by the research evidence. The Manifesto challenges practitioners, as applied scientists, to ensure that their decisions and actions are consistent with the best, most current empirical evidence.

Now, let’s consider some of the specific questions on Issue 2. Several questioners raised a common concern: “How should one proceed in the absence of data?” “How do you deal with the demand for services where there are no established remedies?” Don’t we expect professionals to proceed effectively in the absence of specific protocols and “to use good judgment and manage uncertainty better than nonprofessionals?” Shouldn’t psychologists be allowed to employ “routine and accepted practices” in the absence of empirical evidence? Shouldn’t psychologists “reserve the right to practice some things of unproven validity until the science catches up?”

On one hand, if there is no valid scientific evidence telling us what to do or what to expect in a particular case, then we cannot pretend to have special knowledge; we must admit that we are unlikely to do better than chance or base rates at predicting events or intervening in people’s lives. To be ethical and responsible, under these circumstances, we should not offer “professional” services. On the other hand, if valid

scientific evidence supports the use of a particular assessment or treatment method in a given case, then we not only are free to use this method (assuming that we are competent to use it), but we are obliged to choose this method over other, less effective or untested methods. Some psychological practitioners may not like this answer. It clearly restricts their freedom to act as they wish or to rely on their clinical judgment or experience; however, if psychologists are to be held accountable for their actions, this seems to be the only reasonable reply. One questioner put the matter succinctly: “Would the issue . . . even come up if we were talking about medicine or engineering?”

Most of the remaining questions on Issue 2 focused on how to handle all other cases: that is, those that fall somewhere between the two extremes of “no relevant empirical evidence” and “clearly relevant empirical evidence of incremental validity.” At last, we have come to the crux of the issue. Research evidence seldom fits exactly the unique circumstances of each clinical case we face. So what do we do when “research findings are based on narrow patient samples, unrepresentative therapists, etc.?” How do we deal with “patients whose problems are too complex” to be covered exactly by the research literature? When using empirically supported methods, how do we handle procedural matters “not addressed in the literature?” When describing the exact nature of our empirically supported services, “how exact is exact?” To answer these questions, let’s return to the earlier discussion of a probabilistic scientific epistemology.

I argued that valid theories and methods have incremental validity: that is, they have been shown empirically to reduce uncertainty when applied to a reasonable sample of controlled test cases. Now, let’s apply these concepts to the decisions and predictions of applied psychological scientists. Consistent with the First Corollary, let’s assume that the practitioners who offer psychological services have good reason to believe, based on evidence from controlled empirical research, that such services have incremental validity over a large sample of cases. But how does the applied scientist decide what degree of confidence is reasonable in a particular case, where $N = 1$?

Most of us already know the answer to this question, but let’s review it. Three main factors affect our level of confidence when generalizing from research to practice. The first is the quality of the evidence. Evidence from a well-designed, carefully controlled, experimental study inspires far more confidence than, say, evidence from a quasi-experimental series of clinical case studies. Suppose, for instance, that a clinician approaches each client as an $N = 1$ study, using an ABAB design to assess the effects of each intervention. This strategy is better than none, but it is not as powerful as a controlled experimental strategy because it cannot rule out as many rival explanations for observed effects (Campbell & Stanley, 1963). Similarly, replicated findings inspire more confidence than findings from a single study; we want to know that the results are reliable. And findings based on rep-

resentative samples are more compelling than findings based on unrepresentative samples (Maher, 1978). Finally, results with low variance inspire increased confidence; the narrower the range of likely outcomes, the better our odds of predicting the outcome in a specific case.

The second factor affecting our level of confidence when generalizing from research to practice is the importance of the evidence—both its strength and its practical implications. Strength is reflected in the distance between means, magnitude of an increment, or effect size. Judgments about the strength of an effect will vary as a function of our choice of outcome variables, our decisions about how to capture these variables, the validity of our measures, and our method of analysis. For example, we may find whopping effect sizes on self-report measures of clients’ satisfaction with a treatment but minuscule effects on behavioral measures of change in symptom level. Particularly in clinical settings, we must attend not only to the quantitative magnitude of observed effects, but also to their practical implications. Suppose, for example, that we compared the effects of an experimental and control treatment in a weight-loss program, and found that the group means for individual subject weights were comparable prior to treatment (both about 400 lbs) but differed significantly following treatment (360 and 410 lbs, respectively). In this case, the weight-loss treatment may have yielded a healthy effect size, but it did not have much effect on the size of the real-world health risks faced by these obese subjects.

The third factor affecting our confidence is the degree of similarity between the research and clinical situations. We touched on this point previously in connection with Issue 1. In the language of generalizability theory (Shavelson, Webb, & Rowley, 1989), any two situations can be compared on an infinite number of *facets*. The degree to which the probability distribution of results from one sample will be predictive of the probability distribution of results from another sample is a function of the degree of similarity between the two samples. Not all facets exert an equivalent influence on the generalizability of results, however; indeed, one aim of generalizability analyses is to determine empirically the differential weights and slopes of the generalization gradients for selected facets. Suppose, for example, that a generalizability study revealed that individual differences in clients’ heights and weights were associated only weakly with the outcomes in a controlled treatment study, whereas differences in verbal intelligence and age were associated strongly with the outcomes. If our research sample and clinical sample differed on height or weight, this should have little influence on our confidence that the research results would generalize; but if our samples differed on IQ or age, this should diminish our confidence. The trick, of course, is to decide which facets might be important, assess the association between these facets and outcomes, assess the degree of similarity between the research and clinical samples on these facets, and then use a weighted formula for these facets to predict the clinical

outcome. Thus, we use empirical information to adjust our confidence in the generalizability of research results.

This is difficult to do in practice, of course, because (a) we know too little about which of the infinite number of possible facets actually are critical to clinical outcomes, and (b) even if we knew which facets to select when comparing the research and clinical settings, the empirical literature seldom provides the necessary information for a generalizability analysis. Reports of clinical research trials typically provide only the most rudimentary demographic information about client samples, even less information about therapists, and virtually no information about background procedural details, such as timing of appointments, etc. Furthermore, research clients typically are preselected to provide a purified sample with a specific diagnosis. This enhances internal validity at the expense of external validity; the research sample is less representative of the “complex” clients found in “real-world” clinical settings. These examples illustrate the main point. The generalizability of research results is a function of the similarities between the research and clinical samples across an infinite number of facets, with some facets being more important than others. Unfortunately, we often do not know which facets are important and have no basis for estimating the slopes of the generalization gradients for the different facets.

So, what are the implications of generalizability issues for the First Corollary? Specifically, how do we deal with the fact that no two clients, therapists, or situations—indeed, no two samples of anything—ever are exactly alike? If we wait around for a perfect match before generalizing, we will wait forever. One questioner put it this way: “Clinical psychology seems to think that evidence-based practice means that practitioners must sit around waiting for ‘a study’ to be done on any and every problem they face,” and that “in the absence of ‘a study’ they are free to do whatever seems like a good idea at the time or whatever it is they do regularly.” Applied scientists always should base their clinical decisions on the best research evidence available, even when it isn’t perfect; that way they’ll always be making the best informed decisions possible under the circumstances. Sometimes, choosing to *do nothing* is the best decision (see the Fourth Corollary, point 3, in the Appendix).

Generalizing from one sample to another *always* requires making simplifying abstractions: attending to certain facets while ignoring many others. We also must realize that the number of facets on which any two samples might be compared is limitless; even if two samples are comparable on every facet of interest, they still are sure to differ on other unknown facets. As Cronbach (1975) observed, “Once we attend to interactions [to explain observed outcomes], we enter a hall of mirrors that extends to infinity. However far we carry our analysis—to third order or fifth order or any other—untested interactions of a still higher order can be envisioned” (p. 119). Not only is there an infinite number of potential attributes (facets) and interactions, but the relation-

ships, as captured by interaction equations, can change dramatically as we decide to include or exclude attributes. In short, there is no complete and final answer to the question of which combination of attributes (facets) ultimately accounts for outcomes. Instead of searching for such illusory absolute truths, we need to be more modest and pragmatic, approaching prediction questions as actuaries, using any probabilistic information we can find to improve the accuracy of our predictions in a particular task.

I have argued that our best hope for achieving incremental validity, as applied scientists, is to use the best empirical information available from controlled research to make nomothetic probabilistic actuarial predictions for a sizable sample of clinical cases, even though we may be taking the cases one at a time. Basing our predictions on the probability distribution for a large sample of similar cases, we mitigate the “noise” created by the multitude of unexplored facets that affect outcomes in unknown ways. Inevitably, the accuracy of our probabilistic predictions will vary considerably from one case to another; nevertheless, this approach will yield the highest overall level of accuracy when computed for a large sample of individual cases. Practitioners often are tempted to override actuarial predictions with their clinical judgments when working with a sample of $N = 1$. The preponderance of the research evidence demonstrates clearly that practitioners should resist this urge and trust the actuarial approach (see Grove & Meehl, 1996).

Even when the research samples and clinical samples differ, or when the research literature provides no direct information on procedural details, there is no reasonable alternative to basing predictions and decisions on the probability distributions from the closest research studies. The more precise the information, the better. If information about conditional probabilities is available, use it; if not, the best predictor for a given case will be the probability distribution from the entire research sample.

One questioner asked, “How do you justify generalizing scientific research findings to applied practice when the data supporting doing so are still quite sparse?” The obvious answer, given this analysis, is that we justify it because we simply have found no better epistemological system for maximizing the accuracy of our clinical decisions. If we had a better alternative, we’d choose it; but there really are only two choices: generalize from the “closest” scientific research evidence or make no claims at all regarding a special ability to predict.

Let’s consider a concrete example. Suppose you were an applied clinical scientist, deciding whether to accept a prospective male client. He is 45 years old, divorced, a high school dropout, and unemployed. His primary symptom is anxiety, predominantly displayed as panic attacks and agoraphobia, but he also has other comorbid symptoms, including depression, a history of alcohol abuse, and many failed interpersonal relationships, among other things. You review the research literature and find good empirical support from ran-

domized clinical trials for the use of a manualized, time-limited, cognitive-behavioral treatment (CBT) for panic disorder (e.g., Barlow, Craske, Cerny, & Klosko, 1989; Brown & Barlow, 1995; Practice Guidelines Coalition, 1999). You note, however, that the research clients tended to be female, younger, better educated, married, from a higher socioeconomic level, and less agoraphobic. They also were screened to exclude anyone with a primary disorder other than anxiety, and some clients may have received free treatment in exchange for participating in the trials. Given these differences, should you generalize from the research and accept the client?

If, in your judgment, the research provides a reasonable level of empirical support for the incremental validity of CBT for such clients, and if you are competent in the CBT method, then you have grounds for accepting the client. Otherwise, you don't. But if you accept the client on this basis, you then are obliged to adhere to the empirically supported CBT protocol. You have no grounds for improvising in the hope of making the protocol even more effective for this client. Adherence to the protocol does not mean that you mechanically read from the manual or that you show robotic insensitivity to the client. It simply means that you make a genuine effort to replicate both the content and spirit of the original protocol. To optimize the outcome in this case, you should make actuarial decisions based on the probability distributions reported in the most relevant controlled studies available. Finally, you must gather outcome data, not only to monitor treatment effects in this one client, but also to allow you to evaluate the overall success of the CBT protocol when it is applied to a sample of similar cases and the results are aggregated.

Some readers will ask, "But what should we do, if anything, about our hypothetical client's other problems—his depression, alcohol abuse, and interpersonal difficulties?" Therapists cannot possibly treat everything at once; the usual solution is to employ a triage strategy, focusing first on the most critical problem and, when that is under control, moving to the next most critical problem. However, for each problem, in turn, the clinician should follow the same basic decision process as outlined for the client's panic disorder.

Other readers might ask, "In this hypothetical case, should the clinician combine CBT with a pharmacological treatment?" Once again, the answer to this question should be based on research evidence (e.g., see Telch, Sloan, & Beckner, 1998). My response to all such questions is that applied scientists must base their management decisions in individual cases on the best research evidence available. This kind of empirically based decision process has been described in considerable detail elsewhere (e.g., Hayes, Barlow, & Nelson-Gray, 1999).

Several questioners expressed doubts about the relevance of controlled clinical trials to the kind of complicated cases more typical in everyday clinical practice. Interestingly, there now is research evidence bearing directly on this very ques-

tion, including evidence with panic-disorder cases like our hypothetical client. Wade, Treat, and Stuart (1998) used a "benchmarking strategy" to test the effectiveness of treating an unselected sample of panic-disorder clients in a community mental health center (CMHC) with the same manualized, time-limited, CBT protocol that had been shown to be effective for treating a carefully selected sample of panic-disorder clients in controlled clinical trials (Barlow et al., 1989; Telch et al., 1993). By comparing the CMHC and research samples on a number of behavioral and psychological benchmark measures, Wade et al. provided compelling evidence that the CBT protocol yielded remarkably equivalent positive outcomes in the two settings, despite the obvious dissimilarities in client characteristics, therapists, geographic regions, etc. These favorable comparisons were maintained at a one-year follow-up (Stuart, Treat, & Wade, in press). Thus, the good news is that the CBT protocol was transportable! Practitioners routinely should collect such benchmark data; but in the meantime, the support for generalizability is encouraging.

From this discussion of the Manifesto's implications for clinical practice, we can identify four competencies that differentiate applied clinical scientists from applied nonscientists and technicians. First, applied scientists must be competent at reviewing, evaluating, interpreting, and guiding their actions by the relevant scientific research literature. Second, they must be competent at quantitative reasoning, measurement, and probabilistic inference. Third, they must be competent at hypothesis testing, research design, the systematic and unbiased collection of outcome data on clinical cases, and the evaluation of these data relative to benchmarks drawn from the research literature. Fourth, applied scientists invariably must become specialists in specific theories and methods for specific clinical problems. They cannot possibly be experts in everything; yet they most certainly must be experts in everything they do. Indeed, specialization helps to ensure both the highest level of professional competence and the highest success rate (see Gawande, 1998). In addition, specialization may be the only plausible way for most applied scientists to collect a sufficiently large and homogeneous sample of clinical cases (one at a time) to evaluate the overall accuracy of their predictions and the effectiveness of their interventions, relative to appropriate scientific benchmarks. These four competencies, which are the sine qua non of applied clinical science, are acquired primarily through advanced training in psychological clinical science. This leads us logically to Issue 3, which focuses on the Manifesto's implications for doctoral training.

Issue 3: Questions on Clinical Training

- The Manifesto's Second Corollary asserts that "the primary and overriding objective of doctoral training programs in clinical psychology must be to produce the most competent clinical scientists possible." But what if

- “scientist” jobs are not available? What other roles are there for doctoral level clinical psychologists?
- Isn't it legitimate to train scientist-administrators and scientist-program evaluators as well as scientist-scientists?
 - How do you account for the apparent failure of the Boulder Model? Why don't most doctoral-level clinical psychologists trained under this model actually function as scientists, despite high standards, national conferences, accreditation procedures, etc.?
 - The Manifesto recommends diversity in training (“let a thousand flowers grow”) and an empirical evaluation of training outcomes. Why not allow the same diversity in applied settings? Isn't that a double standard?
 - Current accreditation criteria are too specific and rigid, putting strong pressure on programs to require all students to take courses in psychotherapy, individual assessment, etc. Shouldn't the criteria be more flexible?
 - Why aren't doctoral-level psychologists, with behavioral training, preferred over other mental health professionals, such as social workers, nurses, substance abuse counselors, pastoral counselors, etc.?
 - Shouldn't practitioners with better science training come to dominate the mental health care delivery market, forcing nonscience practitioners to retool empirically? Why hasn't good practice driven out bad? Maybe psychotherapy is like pizza: When it's good, it's really good; when it's bad, it's still pretty good!

Reflections on Issue 3: Training Clinical Scientists

A scientific epistemology does not seem to come naturally. If it did, most people would approach life as scientists, and there would be less need for doctoral training programs. Indeed, one of the strengths of a scientific epistemology is that it protects us from ourselves, from the common flaws in our natural patterns of thought. In this respect, scientific thinking, by definition, is counterintuitive. The special qualities that characterize scientific thinking (e.g., tough-minded skepticism, logical discipline, tolerance of ambiguity, persistence, curiosity, and quantitative reasoning) typically must be learned, nurtured, and refined. This is the job of education, in general, and of doctoral-level education and training, in particular.

The Cardinal Principle asserts that all clinical psychologists, regardless of their particular activity context, should function within a scientific epistemology. Thus, all clinical psychologists need to be trained to think and act as scientists. For this reason, the primary and overriding purpose of graduate-level training in clinical psychology is to train clinical scientists. The Second Corollary goes further, however. Training programs cannot be satisfied merely to offer scientific training and hope that it “takes”; they must evaluate their training and hold themselves accountable for the results. Do their graduates actually function as competent clinical scientists? The graduates need not be basic researchers in research

institutes or academic settings to be counted as competent clinical scientists. As one questioner suggested, successful graduates might be scientist-administrators, program designers and evaluators, educators and trainers, policy researchers and advisors, or clinical practitioners and supervisors. According to the two-dimensional model presented earlier, clinical scientists are not defined by their activity context but by their epistemological approach within their chosen activity context.

One questioner asked why Boulder Model training programs in clinical psychology have such a poor record by this criterion, with too many clinical graduates engaging in activities inconsistent with a scientific epistemology. Either these graduates never learned to think as scientists in the first place or they discarded their scientific training in favor of some alternative after graduation. In either case, training programs must accept some responsibility. Granted, faculties have little influence over students once they have graduated; but it would appear that faculties have failed to impress on such errant students, while they were in the program, the value and importance of adhering to a scientific epistemology.

One likely reason for such disappointing results is that clinical psychology, as a field, has not been sufficiently clear about its training goals. It has tried to be all things to all people. Too often, the various training goals have been in conflict. For example, how can we succeed in training students to be hard-headed skeptics at the same time that we are training them to administer tests and treatments that lack empirical support? At the very birth of clinical psychology, Woodworth (1937) warned against the inherent dangers of training models with a heavy emphasis on professional practice. Unfortunately, history has shown that his concerns were justified (Sechrest, 1992). It is time to get our priorities straight. The first order of business in clinical training must be to ensure that our graduates will function as scientists, in whatever roles they fill.

There is plenty of evidence that doctoral programs successfully arm students with scientific facts and technical skills. Students routinely read and critique the literature, pass examinations, write scholarly papers, and demonstrate technical facility with statistical and research methods. Some of these students go on to careers as clinical scientists, but others do not. No one knows for sure what makes the difference. At present, we can demonstrate that several models of graduate training in clinical psychology are better than other models at preparing students to function as clinical scientists, but we cannot yet say that among the better models one clearly is superior to any other. Therefore, accreditation standards and other criteria for the evaluation of training programs must not be too rigid. The issue of how best to train clinical scientists remains an open question. This is why the Manifesto accepts diversity in science training models. To prescribe a standard model for all scientific training at this point, imposing it on all clinical programs, would be premature and unwarranted. To discover how best to train clinical scientists,

we must allow several training models to coexist and evaluate their outcomes. Once we discover a model that clearly is superior, then we can standardize.

This does not mean that in the meantime “anything goes” in graduate training. Far from it. At a minimum, graduate programs must satisfy criteria parallel to those outlined in the First Corollary. Specifically, they should (a) describe in detail their philosophy, curriculum, faculty, students, resources, etc.; (b) describe the aims and likely results of this training; (c) document their record of achieving the intended results of training (e.g., provide records of students’ progress, placements, activities, and achievements); and (d) expose themselves to public scrutiny (e.g., internal and external reviews, accreditation evaluations, etc.), thereby promoting truth in advertising, program integrity, and quality control. Programs not only should track the progress of students in the program, but also should continue to track graduates’ career paths afterward, always looking for ways to improve their success rates at producing competent clinical scientists.

Someone asked, “Why not allow the same diversity in applied settings? Isn’t that a double standard?” No, it is not. In the applied and training settings alike, a scientific epistemology imposes the same constraints on the process and methods by which we evaluate competing models. Diversity and creativity are encouraged in both contexts. But within each context, we are obliged to evaluate the competing models by employing the best scientific methods available: that is, the methods that will yield the most valid interpretations and inferences. Because it is possible to conduct randomized clinical trials in treatment settings, the experimental method is the optimal choice for evaluating clinical assessments and interventions. Because it is not feasible to conduct randomized trials in graduate training (e.g., we cannot assign students randomly to training programs), we must rely on quasi-experimental methods to evaluate training methods (Campbell, 1969; Campbell & Stanley, 1963). In both contexts, however, our decisions must be based on the empirical evidence. And if the evidence indicates that a particular model is superior, we are obliged to employ that model consistently, at least until a better model is found.

Someone else asked why clinical psychologists with behavioral training are not preferred over other mental health providers, such as social workers, nurses, and counselors. Approaching this question as clinical scientists, our answer must be logical, based on the evidence, not emotional, based on professional affiliation or personal gain. The available research on the differential effectiveness of various mental health providers makes it difficult to argue that clinical psychologists deserve to be preferred over other, less expensive mental health professionals (e.g., Berman & Norton, 1985; Clay, 1998). Until recently, clinical psychologists did not need to worry about such comparisons; there was enough fee-for-service business for everyone. But with the advent of managed care, with its new emphasis on cost-effectiveness and accountability, health care administrators now are asking hard questions about the “bottom line.” Psychologists must

justify their existence. Perhaps, for the first time, “good practice” (defined as cost-effective practice) may be driving out “bad practice.” If clinical psychologists could demonstrate that they were more effective or less expensive than other mental health specialists, they might become the preferred providers. But this seems unlikely. The more we know about effective clinical assessments and treatments, the more we realize that these should be actuarially based and algorithmic, following a standard protocol. This means that these services probably could be administered most cost-effectively by specially trained and supervised technicians.⁴

If we cannot demonstrate convincingly that doctoral training in clinical psychology makes us more effective practitioners than other mental health providers, then what unique advantage, if any, does our training give us? At its best, doctoral training in clinical psychology prepares us to function as clinical scientists in ways, and at a level of competence, unmatched by other mental health disciplines. This training prepares us to conduct basic research, of course. But it also prepares us to design and evaluate assessments and interventions; to contribute to public policy decisions; and to train, supervise, and manage mental health service providers. Our strength in clinical science is the foundation on which we should build our future. We should dedicate our training efforts to producing the most competent clinical scientists possible.

Issue 4: Questions on the Manifesto’s Impact

- What is the Manifesto’s likely impact? Will therapists heed it? Will it strengthen or weaken interest in research? Might it alienate its intended audience?
- Is the Manifesto more than an aspiration? Should there be an FDA-like seal of approval? Do you envision “guidelines” or “ultimatums?”
- Why hasn’t SSCP adopted the Manifesto officially as its policy statement or linked it to the mission statement?
- Does the Manifesto apply with different force to different areas of psychology?
- Why not write a Manifesto for applied researchers aimed at making their products more useful to practitioners? For example, research should be more relevant to existing models of practice, should employ more representative samples of clients and therapists, should include suitable follow-up periods, etc.

Reflections on Issue 4: Implications for Action

The Manifesto was a call to action. It urged scientifically oriented clinical psychologists, such as the members of SSCP and AAAPP, to take a more responsible and active role in articulating and promoting the simple scientific and ethical

⁴Peterson (1996) observed that the United States is “the only country in the world that has a large, doctoral level profession of psychology” (p. 1). Our parochial view sometimes leads us to believe that our way of doing things is the only right way, but that may not be so. Indeed, a questioner from Canada wondered what all the fuss was about; their’s is a very different view.

principles outlined in the Manifesto. Clinical scientists needed to become more active not only because this was the right thing to do, but also because it was in their self-interest to do so. In 1971, Rotter warned that, if clinical psychologists did not act more responsibly by actively and explicitly evaluating the validity of their methods and assuring the quality of their services, they would “find themselves restrained from outside . . . as a result of their own failure to do what ethical and scientific considerations require” (p. 2). The Manifesto echoed this warning loudly.

The Manifesto proposed no specific plan of action. It did not call for an FDA-like stamp of approval, for the establishment of an FPA (Federal Psychotherapeutic Administration), for external regulations or guidelines, or for the issuing of “ultimatums.” It simply said that either clinical psychologists would put their house in order or outsiders would do it for them. And it warned that the time for action was running out.

Now, nearly ten years later, health care reform (official and unofficial) has turned these warnings into stark realities. Psychologists now must deal with these realities, whether they want to or not. It no longer is a question of what psychologists think of the Manifesto—whether they will heed it, are alienated by it, are likely to adopt it as an official policy statement, etc. In today’s health care system, accountability no longer is “optional.” Mental health practitioners are in a game of musical chairs, and the music has stopped. Everyone is scrambling to find a secure place within the changing health care system.

Consider all the changes (McFall, in press) in terminology (“behavioral health” is the buzzword), compensation (“fee-for-service” is becoming a thing of the past), reimbursement criteria (services must be “cost-effective” and “medically necessary”), case management (services must be “preapproved” and “time-limited”), and accountability (“utilization review” focuses on bottom-line results). Such changes inevitably are creating a push toward standardization and an industry-wide proliferation of practice guidelines and standards of care (e.g., Hayes, Follette, Dawes, & Grady, 1995; Practice Guidelines Coalition, 1999). Meanwhile, many practicing clinical psychologists are unable to sustain “business as usual” (Murphy, DeBernardo, & Shoemaker, 1998; Phelps, Eisman, & Kohout, 1998; Rothbaum, Bernstein, Haller, Phelps, & Kohout, 1998). System administrators increasingly are hiring MSW therapists rather than doctoral-level clinical psychologists because social workers are “more cost-effective” (Thyer, 1999).⁵ Even as some psychology training programs continue to turn out large numbers of practitioners, the labor market for practicing clinical psy-

chologists never has looked more bleak (Clay, 1998; Murray, 1999). In short, it is too late for abstract warnings about likely future changes. Concrete changes are here, and more are in the pipeline.

Regardless of how you feel about such changes, you cannot afford to ignore their implications for the future of psychology. These implications merely reinforce the basic messages of the original Manifesto. If the subdiscipline of clinical psychology is to remain viable, it must return to its foundations and recommit unequivocally to a scientific epistemology. It must strive to ensure that all psychological activities—in all areas, not just clinical—meet the highest ethical and scientific standards. And it must refocus its training efforts on the primary goal of producing the most competent clinical scientists possible. In light of all of the changes that have taken place in psychology since the Manifesto was published, these three simple points seem more self-evident and compelling than ever.

Clearly, we are entering a new era in clinical science. Many of the old distinctions within psychology are becoming much less important. We are discovering that the most significant advances in psychological science often come from a hybrid crossing of the various areas of psychology and beyond. To solve such deep mysteries as the of origins of schizophrenia, depression, or aggression, we will need to draw on the best theories and methods we can find, regardless of their source. This means that clinical scientists must look beyond their traditional boundaries and develop expertise across multiple areas, such as cognitive, developmental, neuroscience, quantitative, social, etc. This also means that clinical scientists must question and revise old assumptions. For example, what are the most appropriate professional roles in this new era? How should training programs be redesigned to serve these roles? And how can we restore and strengthen the scientific integrity of the field?

Fortunately, many psychologists have been working actively on a number of fronts to advance clinical science. These activities were independent of the Manifesto, of course, but they are united by their common values, vision, mission, commitment, and sense of urgency. I will cite just a few of these at the risk of offending those I leave out. The American Psychological Society was founded in 1988 with the explicit and exclusive mission of advancing psychological science. The American Psychological Association’s Committee on Accreditation has undergone a thorough reorganization and now recognizes clinical science training as one of the legitimate models for accreditation; this has given doctoral programs increased flexibility to expand their scientific training. Division 12 (Clinical) of APA has organized task forces to review the research literature in the areas of assessment and treatment and to publish reports (e.g., Chambless, 1995; Chambless et al., 1996). SSCP has established an electronic network (SSCPnet) for the exchange of views and information among its members on topics relevant (or not) to clinical science. A number of specialized research societies,

⁵Unfortunately, managed-care companies often seem to focus more on the “cost” side than the “effectiveness” side of this equation. But psychologists contribute to this bias to the extent that they resist efforts to develop strict standards of care, such as those advocated in the Manifesto, which would give the managed-care companies the necessary criteria for distinguishing between effective and ineffective (valid and invalid) services and procedures.

such as the Society for Research in Psychopathology, have worked to advance basic scientific research, even as battles over applied issues in clinical psychology were being fought all around them. The Association for the Advancement of Behavior Therapy was a pioneer group in demonstrating how to carry out meaningful basic research on clinical theories and techniques. AAAPP has taken a leading role on several initiatives relevant to clinical science, including prescription privileges for psychologists, standards of practice in psychology, and scientific graduate training in applied and preventive psychology. The Academy of Psychological Clinical Science, founded in 1995, is a coalition of graduate programs (from academic settings and internship settings) committed to training clinical scientists. Each member program went through an application and peer-review process, in which its aims and achievements in scientific training were examined. The organization's primary mission is to promote and enhance the training of clinical scientists. One new organization deserves mention. The Practice Guidelines Coalition, with representatives from all major players in the behavioral health care field (psychiatry, psychology, social work, nursing, consumers, insurance companies, government agencies, etc.), is assembling panels of experts to write guidelines, based on the empirical research evidence, for the treatment of specific behavioral health problems. As this list demonstrates, there has been an explosion of activity on behalf of clinical science over the last decade. We still have a long way to go, but we have made encouraging progress.

A Personal Reflection

The Manifesto elicited strong feelings, pro or con, in many readers. Being passionate about the issues myself, I can appreciate such reactions. The pro-con split reflects an underlying fault line that has separated clinical psychologists into two main camps over many years. In this respect, the Manifesto was not a source of divisiveness; it merely exposed an existing schism. Some of the negative reactions I received, however, had a different, haunting quality, with undertones of personal insult, injury, or betrayal. These reactions have prompted me to close this article by reflecting on the Manifesto's implications at a more personal, human level.

Having been trained as a clinical psychologist, I am not insensitive to the fact that some individuals (therapists, assessors, expert witnesses, supervisors, professors, students, investigators, authors) were offended when I publicly criticized their favorite professional activities for having little or no empirical support, especially when I said that unsupported activities are inconsistent with the highest scientific and ethical principles. I also am not insensitive to the fact that the Manifesto's recommendations, if taken seriously, might force some psychologists out of work. Furthermore, I am not unaware of the fact that the Manifesto's recommendations, if adopted, might undermine some of psychology's most cherished institutions (professional associations, doctoral train-

ing programs, mental health systems). Finally, I am not surprised that some individuals resented it when the value of their front-line effort to help others was challenged by me, an ivory tower researcher. I can see how someone could take the Manifesto personally. But in my own defense, and in an effort to reach out to everyone I've offended, I'd like to explain what I was thinking when I said such insensitive and potentially inflammatory things.

Let me start by telling a true story. When Richard Nixon was president, his attorney general, John Mitchell, appeared as a guest on Johnny Carson's late-night TV show to announce a new government initiative for research on drug addiction. Mitchell explained that the existing funds for drug research were spread across numerous government agencies and programs. Under the new plan, all of the existing funds would be consolidated in one research program in one agency. Then, amazingly, Mitchell announced in detail just what the government's new research program on drug addiction would find!

Mitchell (a lawyer) clearly did not understand how scientific research works. Researchers simply cannot know for sure what they will find until they get there. That is why they do the research. If they knew the outcome in advance, it might make good propaganda, but it would not make good science. Good scientific research reduces uncertainty; it has incremental validity.

If clinical psychologists genuinely care about finding scientific answers to the important puzzles of mental health, then they simply cannot behave like John Mitchell, committing themselves in advance to certain "truths" even before the data are collected and analyzed. And if clinical psychologists genuinely care about helping clients, then they cannot claim that they know how best to assess and treat these clients even before they have collected, examined, and evaluated the research evidence. I was dismayed by John Mitchell's claim to know things that he could not possibly know, given the available evidence. But I find it especially distressing and embarrassing when doctoral-level psychologists—who definitely should know better, thanks to their training—claim implicitly or explicitly to know things that they cannot possibly know. Unhappily, too much of what I see in contemporary clinical psychology fits this description. My distress over this situation is compounded by my awareness that it need not be this way. There now exists a large and growing body of scientific research that provides a solid foundation for a wide range of valid claims and productive activities in clinical psychology. As a matter of conscience, I must speak out when I see such discrepancies, especially in my own field.

It was not my intention to attack or offend individuals in the Manifesto. Rather, it was my intention to declare that clinical psychology had a nude emperor in its midst and to urge psychologists to work actively to find an enduring remedy for this embarrassing and unacceptable situation. I outlined a principle and two corollaries that might help guide psychology's search for solutions. The Manifesto did not

cause clinical psychology's risky state of UV (unvalidated) overexposure; on the contrary, the Manifesto focused on what should be done to reduce and guard against such exposure. If, in the process, I offended individuals, I regret it.

I hope that readers find these elaborate reflections on the simple Manifesto more illuminating than irritating, more in-

spiring than insulting. In the end, however, I hope that we do not allow any remaining differences among us to cause us to lose sight of the all-important values and aims that we share: pursuing truth, advancing knowledge in psychological clinical science, and above all, improving the psychological health and well-being of our fellow human beings.

REFERENCES

- American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist*, *47*, 1597–1611.
- Barlow, D. H., Craske, M. G., Cerny, J. A., & Klosko, J. S. (1989). Behavioral treatment of panic disorder. *Behavior Therapy*, *20*, 261–282.
- Brown, T. A., & Barlow, D. H. (1995). Long-term outcome in cognitive-behavioral treatment of panic disorder: Clinical predictors and alternative strategies for assessment. *Journal of Consulting and Clinical Psychology*, *63*, 754–765.
- Berman, J. S., & Norton, N. C. (1985). Does professional training make a therapist more effective? *Psychological Bulletin*, *98*, 401–407.
- Boorstin, D. J. (1983). *The discoverers: A history of man's search to know his world and himself*. New York: Random House.
- Campbell, D. T. (1969). Reforms as experiments. *American Psychologist*, *24*, 409–429.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand-McNally.
- Chambless, D. L. (Chair), Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically validated psychological treatments: Report and recommendations. *The Clinical Psychologist*, *48*, 3–23.
- Chambless, D. L., Sanderson, W. C., Shoham, V., Johnson, S. B., Pope, K. S., Crits-Christoph, P., Baker, M., Johnson, B., Woody, S. R., Sue, S., Beutler, L., Williams, D. A., & McCurry, S. (1996). An update on empirically validated therapies. *The Clinical Psychologist*, *49*, 5–18.
- Clay, R. (1998, September). Mental health professions vie for position in the next decade. *APA Monitor*, *29*, 20–21.
- Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, *30*, 116–127.
- Feynman, R. P. (1985). "Surely you're joking, Mr. Feynman!" New York: W. W. Norton.
- Gawande, A. (1998, March). No mistake. The future of medical care: Machines that act like doctors, and doctors who act like machines. *Newsweek*, 74–81.
- Gigerenzer, G., & Murray, D. J. (1987). *Cognition as intuitive statistics*. Hillsdale, NJ: Erlbaum.
- Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Krüger, L. (1989). *The empire of chance: How probability changed science and everyday life*. Cambridge: Cambridge University Press.
- Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law*, *2*, 293–323.
- Hayes, S. C., Barlow, D. H., & Nelson-Gray, R. O. (1999). *The scientist practitioner: Research and accountability in the age of managed care* (2nd ed.). Boston: Allyn and Bacon.
- Hayes, S. C., Follette, V. M., Dawes, R. M., & Grady, K. E. (Eds.) (1995). *Scientific standards of psychological practice: Issues and recommendations*. Reno, NV: Context Press.
- Maher, B. A. (1978). Stimulus sampling in clinical research: Representative design reviewed. *Journal of Consulting and Clinical Psychology*, *46*, 643–647.
- McFall, R. M. (1991). Manifesto for a science of clinical psychology. *The Clinical Psychologist*, *44*, 75–88.
- McFall, R. M. (1996). Making psychology incorruptible. *Applied & Preventive Psychology*, *5*, 9–15.
- McFall, R. M. (in press). Training for prescriptions v. prescriptions for training: Where are we now? Where should we be? How do we get there? *Journal of Clinical Psychology*.
- McFall, R. M., & McDonel, E. C. (1986). The continuing search for units of analysis in psychology: Beyond persons, situations, and their interactions. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 201–241). New York: Guilford.
- McFall, R. M., & Townsend, J. T. (1998). Foundations of psychological assessment: Implications for cognitive assessment in clinical science. *Psychological Assessment*, *10*, 316–330.
- McFall, R. M., & Treat, T. A. (1999). Quantifying the information value of clinical assessments with signal detection theory. *Annual Review of Psychology*, *50*, 215–241.
- Meehl, P. E. (1959). Some ruminations on the validation of clinical procedures. *Canadian Journal of Psychology*, *13*, 102–128.
- Meehl, P. E. (1971). High school yearbooks: A reply to Schwarz. *Journal of Abnormal Psychology*, *77*, 143–148.
- Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, *46*, 806–834.
- Murphy, M. J., DeBernardo, C. R., & Shoemaker, W. E. (1998). Impact of managed care on independent practice and professional ethics: A survey of independent practitioners. *Professional Psychology: Research and Practice*, *29*, 43–51.
- Murray, B. (1999). Some clinical psychology students' growing debt burden triggers response from training programs. *APA Monitor*, *30*, 16.
- National Academy of Sciences. (1995). *On being a scientist: Responsible conduct in research*. Washington, DC: National Academy Press.
- Peterson, D. R. (1996). Making psychology indispensable. *Applied & Preventive Psychology*, *5*, 1–8.
- Phelps, R., Eisman, E. J., & Kohout, J. (1998). Psychological practice and managed care: Results of the CAPP practitioner survey. *Professional Psychology: Research and Practice*, *29*, 31–36.
- Popper, K. (1962). *Conjectures & refutations*. New York: Basic Books.
- Practice Guidelines Coalition. (1999). *Draft guideline for panic disorder*. Reno, NV: Author.

- Rothbaum, P. A., Bernstein, D. M., Haller, O., Phelps, R., & Kohout, J. (1998). New Jersey psychologists' report on managed mental health care. *Professional Psychology: Research and Practice, 29*, 37–42.
- Rotter, J. B. (1971). On the evaluation of methods of intervening in other people's lives. *The Clinical Psychologist, 24*, 1–2.
- Sechrest, L. (1963). Incremental validity: A recommendation. *Educational and Psychological Measurement, 23*, 153–158.
- Sechrest, L. (1992). The past future of clinical psychology: A reflection on Woodworth (1937). *Journal of Consulting and Clinical Psychology, 60*, 18–23.
- Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.
- Shavelson, R. J., Webb, N. M., & Rowley, G. L. (1989). Generalizability theory. *American Psychologist, 44*, 922–932.
- Stuart, G., Treat, T. A., & Wade, W. A. (in press). Effectiveness of an empirically based treatment for panic disorder delivered in a service clinic setting: One-year follow-up. *Journal of Consulting and Clinical Psychology*.
- Telch, M. J., Lucas, J. A., Schmidt, N. B., Hanna, H. H., Jaimez, T. L., & Lucas, R. A. (1993). Group cognitive-behavioral treatment of panic disorder. *Behaviour Research and Therapy, 31*, 279–287.
- Telch, M. J., Sloan, T., & Beckner, V. (1998). Current status of combined treatments: Implications for prescription authority. In S. C. Hayes & E. M. Heiby (Eds.), *Prescription privileges for psychologists: A critical appraisal* (pp. 163–195). Reno, NV: Contact Press.
- Thyer, B. A. (1999). Clinical behavior analysis and clinical social work: A mutually reinforcing relationship. *The Behavior Analyst, 22*, 17–29.
- Wade, W. A., Treat, T. A., & Stuart, G. L. (1998). Transporting an empirically supported treatment for panic disorder to a service clinic setting: A benchmarking strategy. *Journal of Consulting and Clinical Psychology, 66*, 231–239.
- Woodworth, R. S. (1937). The future of clinical psychology. *Journal of Consulting Psychology, 1*, 4–5.

APPENDIX

Excerpts from "Manifesto for a Science of Clinical Psychology" (McFall, 1991)

Cardinal Principle. Scientific clinical psychology is the only legitimate and acceptable form of clinical psychology.

First Corollary. Psychological services should not be administered to the public (except under strict experimental control) until they have satisfied these four minimal criteria:

1. The exact nature of the service must be described clearly.
2. The claimed benefits of the service must be stated explicitly.
3. These claimed benefits must be validated scientifically.
4. Possible negative side effects that might outweigh any benefits must be ruled out empirically.

Second Corollary. The primary and overriding objective of doctoral training programs in clinical psychology must be to produce the most competent clinical scientists possible.

Excerpts from "Making Psychology Incorruptible" (McFall, 1996)

Third Corollary. A scientific epistemology differentiates science from pseudoscience. According to this epistemology:

1. Skepticism is the appropriate and legitimate stance toward all claims about psychological services.
2. The burden of proof regarding the validity of a psychological service rests squarely with the proponents of that service.
3. Skeptics are not required to prove the negative case. The absence of negative evidence is not equivalent to positive support for the validity of a service.
4. Untested services do not deserve special status; the world is full of untested notions. Skeptics must treat untested services as "invalid" until convinced otherwise by the empirical evidence.

5. Claims about outcomes and theoretical explanations for those outcomes must be tested separately. For example, when evidence shows that a treatment is beneficial, it is a logical fallacy ("affirming the consequent") to conclude from this that the theoretical explanation for this effect also is correct.
6. Results are specific. Positive results cannot be generalized (with high confidence) to untested problems, stimuli, methods, therapists, patients, measures, conditions, etc. Small changes sometimes produce dramatically different results.
7. Decisions based on nomothetic evidence are more valid, on the whole, than idiographic decisions based on clinical intuition and judgment.

Fourth Corollary. The most caring and humane psychological services are those that have been shown empirically to be the most effective, efficient, and safe. Genuine caring requires the highest level of scientific rigor. Anything less, no matter how well intentioned, is likely to be less beneficial for the individuals being served.

1. Scientific rigor requires that assessment and treatment protocols be specified in as much detail as possible, validated as specified in the protocol, followed faithfully in clinical applications, and monitored objectively—both in administration and results—in individual cases.
2. The most compassionate procedure for choosing a protocol is one that promotes a fully informed choice, based primarily on a careful review of the scientific evidence and secondarily on a conservative appraisal of the local circumstances.
3. The overriding concern of service providers must be to avoid doing harm or making matters worse. Withholding untested and unproven services usually is the most caring and responsible choice.