

2015

DEVELOPMENT AND PSYCHOMETRIC EVALUATION OF A CLINICAL BELIEFS QUESTIONNAIRE FOR LICENSED PSYCHOLOGISTS

Joseph F. Meyer III
University of Rhode Island, joe.meyer518@gmail.com

Follow this and additional works at: https://digitalcommons.uri.edu/oa_diss

Terms of Use

All rights reserved under copyright.

Recommended Citation

Meyer, Joseph F. III, "DEVELOPMENT AND PSYCHOMETRIC EVALUATION OF A CLINICAL BELIEFS QUESTIONNAIRE FOR LICENSED PSYCHOLOGISTS" (2015). *Open Access Dissertations*. Paper 339.
https://digitalcommons.uri.edu/oa_diss/339

This Dissertation is brought to you by the University of Rhode Island. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu. For permission to reuse copyrighted content, contact the author directly.

DEVELOPMENT AND PSYCHOMETRIC EVALUATION OF A CLINICAL
BELIEFS QUESTIONNAIRE FOR LICENSED PSYCHOLOGISTS

BY

JOSEPH F. MEYER III

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

CLINICAL PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND

2015

DOCTOR OF PHILOSOPHY DISSERTATION

OF

JOSEPH F. MEYER III

APPROVED:

Dissertation Committee:

Major Professor: David Faust, Ph.D.

Robert Laforge, Sc.D.

William Krieger, Ph.D.

Nasser H. Zawia, Ph.D.

DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2015

ABSTRACT

In recent decades, a specific class of dubious clinical practices has been labeled *pseudoscientific* and highlighted as a growing area of concern in psychology. Experts have identified numerous examples of pseudoscientific treatments, which are troubling for various ethical reasons. In light of the absence of research investigating the nature of professional beliefs and knowledge associated with scientifically substantiated and unsubstantiated clinical interventions, the primary objective of this study was to develop a questionnaire (viz., the Clinical Attitudes and Knowledge Questionnaire, or CAKQ) to appraise specific clinical knowledge domains and attitudes toward science among licensed, doctoral-level practitioners of clinical psychology. This aim was pursued through generating items designed to detect the presence of knowledge pertaining to (a) legitimate and questionable treatment techniques used in contemporary clinical practice; (b) general clinical psychology research (e.g., controversies relevant to applied practice); and (c) clinical judgment and decision-making procedures. A preliminary scale consisting of items addressing practitioner attitudes toward science in clinical psychology was also created. A secondary study aim was to ascertain whether psychologists' professed knowledge varied in relation to years involved in clinical practice. Two thousand randomly selected licensed psychologists in New England engaged in clinical practice were invited to participate in the study, and the final sample size was 324 participants. Statistical analyses indicated that the initial hypotheses were partially supported. The hypothesis that a four-component solution would best summarize the CAKQ data was not supported by principal components analysis results. However, the hypothesized relationship

between clinical knowledge and critical thinking skills was partially supported.

Consistent with expectations, a lower reliance on intuitive thinking styles was associated with greater clinical knowledge. Finally, the hypothesis that total number of years of clinical experience would not predict higher clinical knowledge scores was also upheld. Study limitations and future research directions were discussed.

ACKNOWLEDGEMENTS

The author would like to express his sincere gratitude to his major professor, David Faust, for his intellectual nurturance and guidance, invaluable corrective feedback, good-humored tolerance of his professional growing pains, and kind moral support during the vicissitudes of graduate school and fatherhood. Special thanks is also due to members of the author's dissertation committee for their much-appreciated guidance, and to Bob Laforge and Will Krieger for their enthusiastic support of the author's evolving program of scholarly work over the last several years. The author also conveys his heartfelt appreciation for the warm encouragement and forbearance of his family and close friends throughout the many late evenings of dissertation writing.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	v
LIST OF TABLES	vi
LIST OF FIGURES	vii
CHAPTER 1	1
INTRODUCTION	1
CHAPTER 2	7
REVIEW OF LITERATURE	7
CHAPTER 3	68
METHODOLOGY	68
CHAPTER 4	79
RESULTS	79
CHAPTER 5	86
DISCUSSION	86
FOOTNOTES	100
APPENDIX	125
CLINICIAN SURVEY	125
BIBLIOGRAPHY	133

LIST OF TABLES

TABLE	PAGE
Table 1. Demographic and Professional Characteristics of Participants ($N=324$)...	108
Table 2. Clinical Attitudes and Knowledge Questionnaire Item Responses	111
Table 3. Scale Means, Standard Deviations, and Intercorrelations	120
Table 4. Means, Standard Deviations, Rotated Factor Loadings, and Communalities for the Rational-Experiential Inventory Items	121
Table 5. Multiple Regression Analysis Summary for Rational-Experiential Inventory Total Scores and Critical Thinking Questionnaire Subscale Scores Predicting Clinical Attitudes and Knowledge Questionnaire Total Scores	122

LIST OF FIGURES

FIGURE	PAGE
Figure 1. Pareto Bar Graph of “Not Familiar” Responses to Clinical Attitudes and Knowledge Questionnaire Items	123

CHAPTER 1

INTRODUCTION

Economist Alan Blinder (1987) eloquently argued that although having one's heart in the right place, or "softheartedness," was essential to guiding humane approaches to economic policies, a reticence to steer the development and implementation of such policies with well-informed critical thinking, or "hardheadedness," may result in undesirable or even disastrous consequences. Likewise, in clinical psychology, effective practitioners should be capable of not only exhibiting warmth and empathy during therapeutic interactions with clients, but also be willing to keep strictly intuitive and emotionally-driven preferences at bay while scrutinizing the evidentiary value of assessments and treatments to be considered and implemented. Unfortunately, "hardheaded" critical thinking appears to be underutilized (and possibly underappreciated) among a considerable percentage of practitioners within the field (see Gaudiano, Brown, & Miller, 2011; 2012; Sharp, Herbert, & Redding, 2008), which is concerning given the potentially baneful ramifications for vulnerable mental health clients.

A possible consequence of tenuous critical thinking habits combined with insufficient familiarity with peer-reviewed research (although additional factors are likely involved) is the proliferation of suboptimal clinical approaches (Beyerstein, 2001; Lilienfeld, Ruscio, & Lynn, 2008; Lohr, Fowler, & Lilienfeld, 2002, summer). A specific class of questionable clinical practices has been labeled *pseudoscientific* in

recent decades and highlighted as a special area of focus and concern in psychology (Lilienfeld et al., 2003; Pignotti & Thyer, 2009; Still & Dryden, 2004). Despite difficulties formally defining pseudoscience, which is largely attributable to the complexities of the enduring demarcation problem in philosophy of science (Derksen, 1993; Lakatos, 1974; Laudan, 1983; Pigliucci & Boudry, 2013; Resnik, 2000), there is some consensus on specific hallmarks or warning signs distinguishing pseudoscience from science. These include: (a) a disproportionate focus on hypothesis confirmation at the expense of adequate testing and refutation; (b) attempts to shield the fundamental principles (or *hard core*) of a research program from falsification through an ongoing invocation of auxiliary hypotheses, which Lakatos (1970) termed a *degenerative research program* when taken to extremes; (c) evasion of the peer review process; (d) the use of obscurantist jargon to create a superficial veneer of scientific legitimacy; (e) a consistent lack of self-correction and subsequent stagnation of ideas; (f) shifting the burden of proof from claimants to skeptics (e.g., declaring that the onus lies squarely on the critics of an approach to adduce evidence against it); (g) an absence of specified conditions under which claims do not hold (i.e., the delineation of boundary conditions); (h) lack of connectivity with related areas of scientific knowledge; and (i) an overemphasis on personal anecdotes and testimonials to lend credence to claims (Bunge, 1984; Lilienfeld, 2005, September; Lilienfeld et al., 2003; see also Ruscio, 2005). Thus, pseudoscience can be conceptualized as “nonscience masquerading as genuine science” (Lilienfeld, 2010, p. 286) and often contains expansive or extraordinary claims that lack essential supportive evidence, contradict well-established scientific findings, and/or overstep the boundaries of

current scientific knowledge.

Experts have identified numerous examples of pseudoscientific psychological treatments, which include (but certainly are not limited to), (a) *power therapies*, such as Eye Movement Desensitization and Reprocessing (EMDR) therapy and Thought Field Therapy (TFT) for treating symptoms of trauma (Deville, 2005; Herbert et al., 2000); (b) facilitated communication (FC) for the treatment of autism (Jacobson, Mulick, & Schwartz, 1995); (c) neuro-linguistic programming (NLP) for addressing a variety of psychological and emotional difficulties (Witkowski, 2012); (d) rebirthing therapy for alleviating supposed lingering psychological trauma attributable to one's birthing experience (Singer & Lalich, 2008); (e) critical incident stress debriefing (CISD) for the early prevention of trauma symptoms (Deville, Gist, & Cotton, 2006); (f) acupuncture for anxiety (Errington-Evans, 2012) and depression (Ernst, Lee, & Choi, 2011); (g) dolphin-assisted therapy (DAT), which involves swimming and interacting with captured dolphins, for treating developmental disabilities (Marino & Lilienfeld, 2007); and (h) hypnosis as a repressed memory recovery technique (Lynn, Loftus, & Lilienfeld, 2008). In addition, a subset of assessment methods have been roundly criticized for possessing poor validity for various purported uses, such as graphology, or the measure of personality traits via handwriting analysis (Dazzi & Pedrabissi, 2009); and the Draw-A-Person Test (ter Laak, de Goede, Aleva, & van Rijswijk, 2005).

Regardless of how we choose to categorize and label ineffective and/or potentially harmful clinical methods, which has been hotly debated in the psychological literature (see McNally, 2003, and multiple spirited responses in the

same issue), there appears to be overall agreement among scholars that such methods are concerning for various ethical reasons (e.g., the violation of general principles and specific standards of the American Psychological Association (APA) Ethics Code [2002], such as therapists taking care to do no harm and allowing research evidence to guide their practices). From another perspective, the dissemination of ineffective clinical methodologies can be viewed as akin to a negative externality in economic terms. That is, there are clear social, emotional, psychological, and financial costs when ineffective (or possibly harmful) interventions are provided to mental health consumers, but clinicians who regularly use them do not necessarily “pay” for these costs. Rather, clinicians may reap financial profits from offering suboptimal services in place of effective ones, which ultimately subtract from (or, in some cases, endanger) client welfare. Furthermore, certain clinicians may never be held accountable for damages inflicted upon their clients unless formal ethics complaints and/or legal charges are filed.

Alongside ethical violations, the proliferation and perceived acceptance of pseudoscience in clinical psychology arguably contributes to unflattering images of the field as partially evidenced by a raft of controversial articles published in 2009 in high-quality scientific journals (e.g., *Nature*; “Psychology: A reality check,” 2009), well-regarded psychological journals (e.g., *Psychological Science in the Public Interest*; Mischel, 2009), and popular news magazines (e.g., *Newsweek*; Begley, 2009, October), many of which openly questioned the legitimacy of applied clinical practice. In the wake of these accusations, the APA responded not by encouraging further research on the veracity of such claims (e.g., by surveying practitioners on preferred

practices) or investigating negative perceptions of the field, but rather by calling for strategies designed to persuade skeptical scholars and the general public to view the field as worthy of science, technology, engineering, and mathematics (STEM) status (APA, 2010, June). The APA's response arguably ignores and obscures a largely unfettered dissemination of questionable clinical activities while elevating the interests of the field and its members over the needs and psychological well-being of consumers of mental health services.

Despite the flurry of political debates and heightened academic focus on distinguishing scientific from non-scientific clinical practices in recent years, relatively few research efforts have examined mental health practitioners' relevant knowledge and attitudes, for example, knowledge about the effectiveness and appropriateness of certain clinical interventions and decision-making methods for specific purposes, knowledge of the general research literature relevant to applied practice, and attitudes toward scientific thinking in clinical work. Obliquely related lines of questionnaire development (e.g., targeting anomalous beliefs inconsistent with scientific thinking) have almost exclusively emphasized assessing for and counteracting the presence of paranormal beliefs among undergraduate students; furthermore, such questionnaires understandably lack items tailored to clinical psychological practices (e.g., McLean & Miller, 2010; Morier & Keeports, 1994; Vyse, 1997; Wesp & Montgomery, 1998). In fact, only three studies to date (viz., Gaudiano et al., 2011; 2012; Sharp et al., 2008) have directly examined clinicians' perspectives on evidence-based and non-evidence-based practices, and these investigations were relatively narrow in scope (e.g., reported theoretical orientations

and circumscribed areas of suboptimal practices, such as TFT).

An exhaustive review of clinical practices that arguably rest on pseudoscience exceeds the scope of the present work. Many such reviews already exist in peer-reviewed journals and book chapters (see previous citations, especially Lilienfeld et al., 2003, and Lilienfeld et al., 2008) and reflect the thoughtful scrutiny of recognized experts in this area of inquiry. Instead, this dissertation attempted to address a current gap in the literature by using survey methodology to examine what practicing clinical psychologists report to know in the context of clinical practice. Collected data were also analyzed for potential associations with pertinent cognitive/information processing variables. The current study thus sought to redress the abovementioned shortcomings in the literature through a preliminary investigation of clinical knowledge profiles among licensed psychologists (e.g., knowledge of the updated treatment and clinical decision-making literature).

CHAPTER 2

REVIEW OF LITERATURE

Clinical Science and Evidence-Based Interventions

“What is this thing called science?”¹ Numerous pithy characterizations of science abound in both popular culture and within the profession of clinical psychology. Renowned physicist Richard Feynman (1985), for example, described scientific inquiry as a bending over backward to refute one’s own hypotheses. In his oft-quoted Caltech commencement address, in which he discussed the critical missing elements of “Cargo Cult Science,” he stated, “It’s a kind of scientific integrity, a principle of scientific thought that corresponds to a kind of utter honesty—a kind of leaning over backwards” (Feynman, 1974, p. 11). Astronomer and astrophysicist Carl Sagan is commonly known for attributing the essence of science to rigorous critical thinking as opposed to a body of knowledge (Sagan & Druyan, 1997). The father of the scientist-practitioner model in clinical psychology, David Shakow (1976), echoed similar sentiments in championing critical thinking over blind faith in particular theories or wishful thinking about desired research outcomes.

When describing the intellectual habits of his academic mentors at University of Minnesota, Paul Meehl (1993) acknowledged philosopher Bertrand Russell’s “dominant passion of the true scientist—the passion not to be fooled and not to fool anybody else” (p. 728). He further noted that if clinical psychologists forego this passion and forget “the two searching questions of positivism: ‘What do you mean?

How do you know?””, they become “little more than be-doctored, well-paid soothsayers” (Meehl, 1993, p. 728). Some psychologists emphasize intellectual honesty combined with a strong resistance to allowing political correctness to influence knowledge, especially out of a desire to feel more “comfortable” about reality (e.g., Hunt, 1999) while others (e.g., Beyerstein, 1997; Lilienfeld, 2010, 2012) maintain that the scientific common ground among allied psychological disciplines and the “hard sciences” (e.g., physics and chemistry) includes protection against confirmation bias combined with a rigorous and conscientious ferreting out of erroneous beliefs and preconceptions.

Despite the aptness and elegance with which luminaries from different fields of study have articulated certain core components of science, it would be impossible to summarize the multifaceted nature of science with a single sentence. Indeed, entire books have been devoted to capturing the essence of science (e.g., Carey, 2011; Chalmers, 1999; Gauch, 2003; Schurz, 2014). Virtually all such books agree that furthering our understanding of nature is made possible by a foundational characteristic of science that distinguishes it from other methods of truth seeking—the *scientific method*. At its most basic level, the scientific method entails observation, testing, and explanation (Carey, 2011). Summarized in more detail (albeit not exhaustively), this process may proceed as follows (cf. Box, Hunter, & Hunter, 1978; Gauch, 2003): Seeking to discover the truth about a phenomenon of interest, a scientist first obtains existing information about the phenomenon (e.g., previous research and observations). This information assists with shaping initial educated guesses about the phenomenon, which ultimately take on tightly defined and testable

hypotheses that subsequently steer the design and methodological features of the study. Pre-established procedures guide the systematic collection of data, which contain truth about reality intermingled with noise (i.e., various kinds of error), the latter being diminished as much as possible via sound research design and conscientious scientific conduct. Next, data bearing on the hypotheses are summarized and appropriately analyzed, and conclusions and inferences are carefully formulated with an eye toward observed results. Finally, future research directions, study limitations, new hypotheses, and/or possible theoretical adjustments are submitted for consideration, and the entire study is summarized in written form to be peer reviewed by the scientific community. If the study stands up to the critical scrutiny of other independent scientists and is replicated by additional investigators, the findings are added to the provisional corpus of scientific knowledge until further revised or later disconfirmed.

An armamentarium of cognitive approaches and logical tools is put to use as the scientific method is applied. This encompasses general problem solving skills, critical thinking, skepticism, inductive logic, deductive logic (or, perhaps more accurately in most practical contexts, abductive reasoning; Fann, 1970), eliminative parsimony, probabilistic thinking, falsification (i.e., attempts at hypothesis disconfirmation), relational reasoning, causal inference, and analogical reasoning (Dunbar & Fugelsang, 2005; Gauch, 2003). Epistemologically, these modes of thinking are typically accompanied by a form of Peircean fallibilism (Peirce, 2011) in contemporary scientific inquiry, which acknowledges the influences of error and uncertainty in rendering knowledge provisional. In line with fallibilist and

probabilistic modes of thinking, inferences drawn in scientific psychology are tainted by layers of uncertainty and are thus subject to the dominion of probability theory (unlike pure mathematics and formal logic). Because determinism in psychology is untenable, psychological theories cannot be sufficiently *proved* in the strong Euclidian sense (Meehl & MacCorquodale, 1991). Rather, psychologists are tethered to more cautious and tentative articulations of research findings. They can declare, for example, that theories are rendered “more likely” or “more credible” based upon the preponderance and comparative weighing of both confirmatory and disconfirming evidence, but declarations of formal proof are erroneous.

Thus, a central objective of scientific research is to arrive at the best possible *approximations* of various aspects of reality (e.g., a correspondence theory of reality or semi-hemi-demi scientific realism; O’Connor, 1975; Irwin, 1988; Russell, 1912) using sound reasoning principles and methodological tools. This epistemological clarification is especially important given that the pursuit of “absolute certainty” is futile “concerning questions of fact” (Peirce, 2011, p. 59). However, scientists do their utmost to allow features of reality (versus strictly socio-cultural constructions of reality) to dictate the content of truth propositions (Boghossian, 2007). In general terms, nature should ultimately control what scientists take to be “correct” answers about it (Irwin, 1988). Thus, truth is conceptualized as an accurate representation of the state of nature insofar as the scientific method can elucidate, although the presence of error and tentativeness of conclusions about worldly phenomena should always be borne in mind.

All of the aforementioned perspectives are compatible with scientific

skepticism, which is rooted in the acknowledgement of a complex world that is unlikely to be fully understood by bounded human understanding (Deville & Lohr, 2008, p. 107). They are also congruent with Donald Campbell's (1974) concept of "evolutionary epistemology" in science, or the ferreting out of erroneous claims over time via the scientific method in a manner akin to natural selection. Ideally, psychological researchers consistently and conscientiously adhere to disciplined cognitive habits (i.e., following the scientific method and utilizing relevant critical thinking skills) when conducting treatment outcome research, and clinical practitioners ideally draw from the same cognitive toolkit when making important decisions about the care of mental health clients. This perspective of scientifically informed researchers and practitioners is consistent with David Shakow's original formulation of the scientist-practitioner model in clinical psychology (Baker & Benjamin, 2000; Cautin, 2008; Shakow, 1969).

The scientist practitioner model. The advent of the scientist-practitioner (or "Boulder") training model occurred during the Boulder Conference on Graduate Education in Clinical Psychology in 1949 (Raimy, 1950). Conference participants agreed that research-informed practice was critical to the maturation and survival of clinical psychology, and the doctor of philosophy (Ph.D.) degree was decreed the legitimate basis for professional licensure (Baker & Benjamin, 2000). From the perspective of Boulder model proponents, practical clinical training must go hand-in-hand with rigorous research training given that these professional activities mutually inform one another in important ways. The ideal scientist-practitioner is thus someone who has the skills and knowledge base not only to make informed and responsible

decisions about clinical care (i.e., by being able to understand, evaluate, and selectively apply research findings), but also to produce and publish novel research to advance the field (Jones & Mehr, 2007). As Belar and Perry (1992) noted, critical thinking is viewed as a centerpiece of this training model according to the National Conference on Scientist-Practitioner Education and Training in 1990.

A second model of graduate clinical training is the Vail or practitioner model, which culminates in the conferral of a doctor of psychology (or Psy.D.) degree. In July of 1973 at a conference in Vail, Colorado, the APA and National Institute of Mental Health officially recognized the Psy.D. model (Donn, Routh, & Lunt, 2000; Mitchell, 1977). Relative to Ph.D. programs, Psy.D. programs are more likely to: (a) focus more heavily on applied clinical work and practice-oriented course material, (b) deemphasize data-driven research endeavors (although incorporating research into practice ideally remains an emphasis), and (c) not require a formal dissertation involving data collection and quantitative data analysis, although many such programs do require scholarly projects and non-data-driven dissertations (e.g., theoretical papers) (Donn et al., 2000; McIlvried et al., 2010). In addition, Psy.D. programs, many of which are housed in non-university, for-profit schools, typically accept more students², offer less financial aid, and place their graduates in more strictly applied clinical settings (vs. academic or research settings) (McIlvried et al., 2010).

These observations are presented here to illustrate key differences between the two dominant training models in contemporary clinical psychology and not to argue for the superiority of one model over another per se, although it may be fair to hypothesize that the typical student may be more likely to encounter stronger scientific

training in a scientist-practitioner or clinical science Ph.D. program compared to the average Psy.D. program. Indeed, as Baker and colleagues (2008, p. 85) suggested, a large number of Ph.D. programs likely only pay “lip service” to scientific foundations, and a small number of Psy.D. programs do offer rigorous science-based curricula. At any rate, differences often highlighted between prototypical scientist-practitioner Ph.D. and practitioner-based Psy.D. programs (i.e., predominantly research- vs. practice-based) unavoidably touch upon a broader concern in contemporary clinical psychology: the rift between science and practice.

The term “scientist-practitioner gap” (Cautin, 2011; Tavris, 2003) is often used to refer to the schism between sound research and effective practice in clinical psychology, although the terminology obviously focuses on the divide between the professionals themselves. Other researchers refer to this same problem more generally as the “science-practice gap” (Lilienfeld, 2013). From one angle, this gap may be viewed as the consequence of researcher-practitioner alienation or isolation (e.g., lack of overlap in professional contact and activity), which may inadvertently contribute to poor clinical care (Teachman et al., 2012). That is, on the one hand, clinical researchers may not have the time, resources, or inclination to focus their professional energies on the dissemination of effective interventions to clinicians. Practitioners, on the other hand, may have limited access to up-to-date research on evidence-based practice, may not have the inclination to seek out advice from clinical scientists, and/or may be reluctant to incorporate scientifically supported techniques into routine practice for various reasons (e.g., antipathy toward science and/or research).

Some scholars, for example, have expressed the belief that most practicing

therapists are hostile toward (and thus ignore) research because they view it as an overly objective approach to a subjective enterprise that attempts to micro-manage their professional activities (Pinsof, Goldsmith, & Latta, 2012, p. 253). A potentially harmful consequence of such antagonistic attitudes (if accurate) may be what *New York Times* columnist Harriet Brown (2013, March 25) referred to as a “dim sum” approach to psychotherapy, or cobbling together various intervention strategies more strongly associated with personal biases and training background than with the up-to-date research literature. Unfortunately, there is little-to-no available formal research at present (insofar as the author is aware) inquiring into therapist attitudes toward research, evidence-based practices, and science in general, thus rendering most of the abovementioned commentary conjectural. Before turning more fully to factors potentially associated with the apparent schism that sometimes exists between research and practice (e.g., pseudoscience, suboptimal clinical decision-making, and other cognitive pitfalls), it is first important to outline the foundations of sound clinical psychological science.

The building blocks of clinical psychological science: RCTs. Conducting sound science in psychology requires careful adherence to scientific method. Researchers are expected to engage in disciplined and systematic observation, hypothesis development and testing, data collection and analysis, problem solving, critical thinking, and the formulation of results and conclusions based on the available data (Deville & Lohr, 2008, pp. 106–108). In 1991, Richard McFall summarized his vision of clinical science within a cardinal principle and two corollaries. His cardinal principle holds that scientifically informed clinical research and practice are the *only*

legitimate forms of clinical psychology. The first corollary admonishes that psychological services should not be publicly disseminated until four conditions have been met, to wit: (a) a clear description of the service, (b) an explicit statement of the hypothesized benefits of the service, (c) scientific validation of the hypothesized benefits, and (d) ruling out potential harmful consequences outweighing service benefits. The second corollary maintains that clinical psychology doctoral programs should prioritize the professional development of “the most competent clinical scientists possible” (p. 79). As McFall (1991) subsequently acknowledges, the most effective training methodologies suited for this objective are not readily identifiable. Nevertheless, a small group of clinical scientists concerned with perceived inadequacies in current graduate school curricula and training created the Psychological Clinical Science Accreditation System (PCSAS) in 1995 in an effort to officially recognize doctoral-level clinical psychology programs and internships offering rigorous training in scientific methods applied to both research and practice (e.g., scientifically-informed assessment and treatment selection and implementation) (Baker, McFall, & Shoham, 2008). However, it is presently unclear what has come of these efforts.

McFall (1991) along with other dedicated psychological scientist (e.g., Klerman, 1990) strongly advocated for the use of randomized controlled trials (RCTs) in clinical research, which is one of the cornerstones of clinical science. RCTs require random assignment of participants to treatment conditions and tightly controlled procedures to ensure (as best as possible) that participants across groups are treated the same with the exception of the intervention component(s) unique to their specific

condition (Bolton, 2008). Stated another way, the crux of controlled clinical trials is to isolate and identify treatment components that exert beneficial effects (e.g., symptom reduction and/or improvements in life functioning) surpassing the effects (or lack thereof) observed in comparison conditions, which may involve no-treatment conditions, waiting-list controls (WLCs), psychological placebos³, or standard practice treatments already deemed effective, which typically comprise the treatment as usual (TAU) category (Bolton, 2008; Kazdin, 2003).

Whereas participants in WLC conditions temporarily receive no treatment at all (i.e., until the conclusion of data collection), those in placebo-control conditions still receive an intervention, albeit subtracting out the critical treatment ingredients hypothesized to confer additional psychological benefits (i.e., over and above no treatment, non-specific therapeutic factors, or TAUs). Psychological placebos assumedly contain all of the generally beneficial yet incidental factors (e.g., treatment credibility, expectancy effects, therapeutic alliance influences, etc.) divorced from specific therapeutic elements (e.g., interoceptive exposure for panic symptoms) hypothesized to confer unique improvements (Herbert & Gaudiano, 2005). Thus, placebo conditions have at least some surface credibility vis-à-vis active treatments (see O'Connor et al., 2007, p. 185, for a clear description of an attention placebo control procedure compared to a CBT condition for treating delusional symptoms).

Of course, psychological placebos, which are vastly different from the inert sugar pills used in pharmacological research (e.g., sugar pills obviously do not engage in human relationships), are not without their share of limitations. These include a lack of perceived treatment credibility to client participants (e.g., transparency of

assignment to placebo-control groups), poor congruence of structural features of psychological placebo-control groups with target treatments (Baskin, Tierney, Minami, & Wampold, 2003), and challenges controlling for and matching expectancy effects between control and treatment conditions (Boot, Simons, Stothart, & Stutts, 2013). In addition, a substantial portion of what has been traditionally viewed as noise or nuisance variance associated with psychological placebos actually consists of legitimate but understudied *non-specific factors* inherent in all forms of mainstream psychotherapy (Norcross, 2011; Omer & London, 1989; Wampold, 2001; Wampold et al., 2010).

Devilly and Lohr (2008) described three key categories of non-specific effects, namely: (a) *common factors*, or recognized ubiquitous features of most treatments imparting therapeutic benefit (e.g., therapist attention, persuasion, expectancy effects, and so forth); (b) *unspecified but active factors*, where “unspecified” refers to palliative features of treatments not explicitly pinpointed as active ingredients (e.g., interpersonal influences embedded in therapeutic procedures and sociocultural contexts); and (c) *factors without specific activity*, or factors inherent in many treatments that more diffusely allay symptoms through ambiguous (non-specific) mechanisms of action (e.g., in the field of medicine, the example of aspirin being used to address disparate physical ailments). Obviously, there is much conceptual overlap among and ambiguity within these broad categories. Furthermore, it is difficult to find methodologically sound studies focusing on concrete, isolatable examples of these factors in contemporary psychotherapy research, although one exception would be the work of Bruce Wampold (2001), who posits a contextual model of psychotherapy (i.e.,

attempting to clearly delineate and systematically study broad classes of non-specific therapeutic actions). And finally, conceptual problems emerge when one recognizes the arbitrariness of distinguishing specific from non-specific factors, which may be dependent on the clinician's theoretical orientation (e.g., accurate empathy would be considered a *specific* factor in humanistic psychotherapies but a *non-specific* factor in CBT) (Herbert & Gaudiano, 2005, p. 897).

Although the effects of non-specific factors can be powerful (see Wampold, 2001) and certainly should not be dismissed, these effects alone are often insufficient for establishing treatment efficacy and differential effectiveness among different forms of psychotherapy (e.g., which treatments work for which conditions, and under what circumstances). Component-controlled efficacy studies assist in clarifying which specific ingredients of a particular novel treatment work with *ceteris paribus* applied to non-specific (e.g., interpersonal and contextual) factors as best as can be achieved in this complex domain of research (Deville & Lohr, 2008). Of note, the importance of examining critical mediators and moderators of therapeutic change likewise should not be underestimated in this context, nor should the value of longitudinal designs for addressing treatment outcome stability and process variables (e.g., specific mechanisms of change⁴ over time) lurking both within and beyond the pre- and post-treatment interval (Laurenceau, Hayes, & Feldman, 2007). Despite their limitations, RCTs comprise a rigorous scientific research design that far exceeds many other clinical decision-making methodologies (e.g., subjective clinical impressions) and remains at the heart of efficacy studies, which will be described next in more detail.

Efficacy and effectiveness research. *Efficacy studies* focus primarily on

symptom reduction and usually are conducted in research clinics staffed by specially trained clinicians and research assistants (Baker et al., 2008; Kazdin, 2003).

Methodological approaches that typically define efficacy research include random assignment, double-blind procedures, placebo or other appropriate treatment comparison groups, standardized assessment (e.g., structured or semi-structured diagnostic interviews), and adherence to pre-specified inclusion and exclusion criteria for research participants (Baker et al., 2008; Chambless & Ollendick, 2001; Campbell & Stanley, 1963). Properly applied, these procedures work together to strengthen the internal validity of efficacy studies, which entails ruling out a complex host of rival alternative explanations (Campbell & Stanley, 1963). In the context of treatment outcome research, specific threats to internal validity include (but are not limited to): (a) the placebo effect (e.g., the combined effects of expectancy, compliance, suggestion, distraction from symptoms, etc.); (b) the self-limiting nature of many psychopathological conditions (i.e., gradual natural recovery irrespective of treatment); (c) spontaneous remission; (d) cyclical symptom variation over time; and (e) the impact of self-serving cognitive biases on self-report data (Beyerstein, 1997).

As Beyerstein (1997) noted, clinical scientists have a professional and ethical responsibility to establish that treatments are safe and effective, the latter almost always being the most challenging in research due to the necessity of ruling out these rival hypotheses. Failure to address and control for these threats via random assignment, the use of placebo-controlled groups, and double blinding often results in an unwarranted attribution of observed recovery (e.g., symptom alleviation or remission) to dubious interventions (see “The Twilight Zone of EMDR” section of this

dissertation for examples of published studies lacking these critical safeguards). However, if these controls are properly implemented, and the magnitude of outcome improvement in the treatment group is meaningfully larger⁵ than observed improvement in the placebo (or no-treatment, WLC, or TAU) group, this suggests treatment efficacy, especially when replicated (Beyerstein, 1997).

In contrast to efficacy studies, *effectiveness studies* are quasi-experimental and are designed to boost external validity (also termed generalizability or ecological representativeness), or the degree to which research findings may be soundly extrapolated to real-world clinical scenarios (although additional cross-validation procedures are critical for establishing external validity; e.g., see Hoepfner et al., 2012). The core question here is the degree to which the intervention works in other applied clinical settings outside of the boundaries of the study. This determination is pursued by relaxing stricter degrees of experimental control (e.g., random assignment) in order to conduct treatments in more naturalistic, representative settings where the usual clinical staff members deliver their typical target interventions to client participants, who are not subjected to as stringent inclusion and exclusion criteria.

Some scholars perceive a tension between efficacy and effectiveness research (i.e., rigor and relevance, respectively) and note that the more tightly controlled a study, the less results will generalize (see Gelso, 1985). In contrast, Baker and colleagues (2008), frame effectiveness research as bridging the gap between research and non-research clinics, and despite some differences in research outcomes comparing efficacy and effectiveness studies, effect size estimates are mostly concordant (Lambert, 2013; Nathan & Gorman, 2007). Ideally, clinical interventions

that are consistently well supported by replicated efficacy and effectiveness results should be broadly disseminated across mental health consumer populations with the objective of serving as many clients in need as possible (McHugh & Barlow, 2010).

Empirically supported treatments. It is a well-replicated finding that, on average, mental health clients receiving psychotherapy evidence more symptom remission and functional gains than those either receiving a psychological placebo-control intervention or not undergoing therapy (see Lambert, 2013). A more contentious and complex question is “what works for whom?” (cf. Roth & Fonagy, 2005), or more explicitly stated, which interventions work best for which psychopathological conditions, and to which populations of mental health clients do these findings apply? In a systematic effort to address the identification of specific treatments backed by sound scientific research, Division 12 (Clinical Psychology) of the APA organized the Committee on Science and Practice (CSP) in 1993 to develop guidelines for defining empirically supported treatments, or ESTs (Task Force, 1995; originally called empirically validated treatments or EVTs; Chambless et al., 1996). Lists of ESTs with accompanying definitions of categories of evidentiary support were subsequently published (Chambless et al., 1996, 1998; see also Chambless & Ollendick, 2001).

According to Division 12 Task Force criteria (see Chambless et al., 1998), treatments deemed *well-established* must be supported by “at least two good between-group design experiments” with evidence of efficacy via either (a) demonstrated beneficial effects over and above placebo or other treatment, or (b) equivalence to a previously established treatment (e.g., TAU) with “adequate sample sizes” (p. 4). An

alternative way for treatments to achieve this status is by demonstrating efficacy in “a large series of single-case design experiments” (conducted by at least two different research teams) involving “good experimental design,” treatment comparison conditions, treatment manuals (or, alternatively, explicitly defined intervention components and steps), and specified sample characteristics (Chambless et al., 1998, p. 4). One step below well-established treatments are *probably efficacious* treatments, which may meet any one of the following three criteria: (a) demonstrated superiority to WLC in at least two separately conducted experiments, (b) meeting all necessary well-established treatment criteria with the exception of independent replication by at least two different research teams, or (c) meeting well-established treatment criteria using a “small series of single-case design experiments” (Chambless et al., 1998, p. 4). Finally, *experimental treatments* are those not yet subjected to the methodological scrutiny the Division 12 Task Force (1995; Chambless & Ollendick, 2001) recommended.

Numerous psychosocial interventions for various psychopathological conditions have met (and surpassed) the Division 12 Task Force’s well-established criteria. For example, multi-component CBT (i.e., with relapse prevention) for tobacco users is an efficacious, effective, disseminable (e.g., easily deliverable by telephone; Chen et al., 2014; Swartz et al., 2005), and cost effective preventative treatment that consistently improves and maintains smoking cessation outcomes across diverse client populations (Hollis et al., 2000; Maciosek et al., 2006; Sheffer et al., 2009) despite its apparent underutilization (Shiffman, Brockwell, Pillitteri, & Glitchell, 2008). Similarly favorable long-term research results support the use of

CBT strategies (especially behavioral interventions) for major depressive disorder (Honyashiki et al., 2014), panic disorder with or without agoraphobia (Otto & Deveney, 2005), bulimia nervosa (Cooper & Shafran, 2008), posttraumatic stress disorder (Foa, Gillihan, & Bryant, 2013), generalized anxiety disorder (Bolognesi, Baldwin, & Ruini, 2014), obsessive-compulsive disorder (Lewin, Wu, McGuire, & Storch, 2014), hypochondriasis (Olatunji et al., 2014), reduced risk of psychosis (Hutton & Taylor, 2014), and various other psychological conditions (see Chambless & Ollendick, 2001, for a comprehensive list of ESTs for specific conditions and Lambert, 2013, for further details). In addition, aside from CBT, interpersonal psychotherapy is also considered a well-established treatment for depression, as is behavioral family therapy for schizophrenia, behavioral marital therapy for marital discord, and brief psychodynamic therapy for geriatric depression (Chambless et al., 1998; Chambless & Ollendick, 2001).

Many vociferous concerns about the concept of ESTs and their use in clinical practice have been raised (e.g., Westen, Novotny, & Thompson-Brenner, 2004). Stewart and colleagues (2012) summarized and offered rejoinders to some of the most commonly encountered (and often fallacious) objections, including supposed lack of generalizability to real-world clinical practice, appeals to clinical intuition and expertise as superior or equal to research findings, the specious yet aggressively pervasive belief in the Dodo Bird verdict (i.e., the unfounded claim of universal treatment equivalence due to the therapeutic alliance, hope, empathy, etc.; see also Hofmann & Lohr, 2010, January), and the charge that ESTs are “unfair” because they overwhelmingly favor CBT or strictly behavioral interventions over other forms of

psychotherapy. More thoughtful and discussion-worthy objections include the observation that placebo-control conditions and WLCs set the evidentiary bar too low (Herbert & Gaudiano, 2005), especially in light of the Edinburgh Revision to the 1964 Declaration of Helsinki (World Medical Association, 2013); and the lack of research-based guidelines to steer idiographic tailoring of ESTs (i.e., scientifically justified specifications for adapting specific approaches to specific conditions for certain kinds of clients) (Holt & Beutler, 2014). The present author agrees with the position that if provisional EST lists indeed turn out to boost the ratio of effective, scientifically supported treatments to largely ineffective, scientifically counterfeit approaches, then these lists are worth compiling and implementing despite their unavoidable shortcomings (Chambless & Ollendick, 2001; Lilienfeld, 2010). However, the objection to the unwieldy impracticality of tailoring many different manualized treatment protocols to various DSM disorder categories is well taken (Wachtel, 2010) and may justify studying principle-focused approaches instead (e.g., the Unified Protocol for Transdiagnostic Treatment of Emotional Disorders, which has received some preliminary research support; Bullis, Fortune, Farchione, & Barlow, 2014).

As an important aside, this discussion of ESTs is unavoidably intertwined with the contemporary healthcare climate in the United States. Specifically, as healthcare delivery systems, expenditures, and economic decision making (esp. of mental health care stakeholders) have continued to change over the past three decades, a growing prioritization of cost-effective mental health interventions has emerged (Baker et al., 2008). Treatments consistently shown to alleviate psychological distress (using assessment tools that quantify symptom severity) in a relatively brief time period over

and above the effects of rival interventions will more likely survive the ever-increasing pressures of managed care (Hayes, Barlow, & Nelson-Gray, 1999). That is, parsimonious interventions (e.g., cognitive-behavioral therapies) supported by scientific research will likely fare better regarding incorporation into the healthcare delivery system and receive insurance reimbursement (Kauth, Sullivan, Cully, & Blevins, 2011). Thus, in the face of dramatically rising costs and service demands in mental and behavioral health care (e.g., see Poisal et al., 2007), it remains incumbent upon doctoral-level practitioners of clinical psychology to offer interventions that are effective, efficacious, cost-effective, and scientifically sound (Baker et al., 2008; Beecham et al., 1997) and to monitor real-life clinical outcomes for accountability and quality improvement purposes (e.g., Hodges & Wotring, 2004). These key quality-of-care criteria are already guiding healthcare coverage decisions as insurance companies and governmental agencies continue to oversee increasingly large swaths of funding in these areas, and this trend is expected to continue well into the future (Baker et al., 2008; Committee on Crossing the Quality Chasm: Adaptation to Mental Health and Addictive Disorders, 2006; Grizzle, Olson, & Motheral, 2000).

Despite the pressing need to implement evidence-based interventions under increasingly stringent managed care guidelines and the large body of research bolstering the manifold benefits ESTs, many psychologists appear unmotivated to disseminate, publicly promote, or offer sufficient training in these interventions (Baker et al., 2008; Stewart et al., 2012). For example, evidence suggests that many clinical psychology graduate students are not receiving adequate training in ESTs, which may partly explain a number of disconcerting survey results among practicing clinicians

(e.g., that 23% of a national sample of 891 psychologists reportedly had never heard of a treatment manual; Addis & Krasnow, 2000). The APA Division 12 Task Force conducted one of the earliest inquiries into this troubling matter, surveying 167 training directors of clinical psychology graduate programs (Ph.D. and Psy.D.) in North America (Crits-Cristoph, Frank, Chambless, Brody, & Karp, 1995). The investigators' (1995) results were unsettling. They found that about one in five doctoral-level, APA-accredited programs did not provide *any* didactic or training in 75% of available research-supported interventions of known efficacy and effectiveness for various disorders; and nearly 50% of training director respondents maintained that competency and training in ESTs was of marginal importance.

Although limited in number, later research studies yielded similarly sobering results (e.g., Horan & Blanchard, 2001, winter; Hays et al., 2002). For example, an exploratory survey of 133 APA-accredited clinical internships found that a mere 28% of the sites provided more than 15 hours of supervision and training in ESTs (e.g., CT, CBT, DBT, and IPT), and furthermore, that 30% of the sites dedicated either minimal (19%) or no (11%) time for EST training opportunities (Hays et al., 2002). Finally, a survey of 172 graduate students drawn from 60 APA-accredited clinical, counseling, and school doctoral psychology programs revealed that close to two-thirds of respondents had never read a single evidence-based treatment related publication (e.g., Task Force articles, treatment manuals, etc.), and approximately 32% had never taken a course covering EST content or research (Karekla, Lundgren, & Forsyth, 2004).

Missing from these surveys is information regarding what exactly many graduate students *are* spending their time learning during their pre-doctoral-training

years in the place of scientifically supported interventions. As Lilienfeld (2010) conjectured, much of the time not dedicated to learning effective treatments may instead be spent learning about non-specific factors (e.g., warmth, empathy, listening, etc.) and/or less-than-optimal intervention techniques. Among the latter lie pseudoscientific treatments, which contribute to the professional marginalization of the field of clinical psychology.

The Contours of Pseudoscience

What is pseudoscience? In the contemporary climate of second-generation managed care and professional accountability, it is incumbent upon mental health professionals to choose their interventions with deference to best evidence (Hayes et al., 1999). To do otherwise, such as on the basis of appeals to novelty or what is emotionally appealing to the therapist or patient, potentially carries harmful consequences for clients and professionals alike. Although there are many widely available evidence-based treatments with sound theoretical underpinnings alongside responsible practitioners who adhere to them, pseudoscientific theories and treatments remain pervasive in the field of clinical psychology (see Lilienfeld et al., 2003).

Far from being a novel phenomenon in the current professional landscape (see Gardner, 1957), pseudoscience⁶ has been repeatedly recognized over time as a major threat to both public welfare and the scientific foundation and integrity of the field (Lilienfeld, 1998, fall). The proliferation of pseudoscience may be partly attributed to anti-science sentiments and the acceleration of commercial marketing of interventions (Olatunji, Parker, & Lohr, 2005–2006, fall/winter). Some psychologists (e.g., past APA president Ronald Fox) apparently deem clinical outcome research superfluous as

evidenced by public pronouncements of the following ilk: “Psychologists do not have to apologize for their treatments. Nor is there an actual need to prove [*sic*] their effectiveness” (Fox, 2000, pp. 1–2). Other factors likely contributing to the proliferation of pseudoscience in psychology include human credulity combined with poor critical thinking skills, which may persist despite education level or intelligence (Shermer, 2002), vulnerability in the face of refractory mental health conditions, self-serving biases (including cognitive dissonance), and false hope, all of which may hasten the suspension of reason (Beyerstein, 1997; Worrall, 1990).

Formal definitions of pseudoscience have proven difficult. As social psychologist Carol Tavris once remarked at an American Psychological Society symposium, “Pseudoscience is like pornography; we can’t define it, but we know it when we see it” (as cited in McNally, 2003, p. 97). However, there is substantial agreement on an interrelated set of general hallmarks or warning flags of pseudoscience (Beyerstein, 1997; Bunge, 1984, fall; Derksen, 1993; Hines, 2003; Lilienfeld et al., 2003; Pratkanis, 1995; Ruscio, 2005; Stanovich, 2003), most of which were outlined in the Introduction of this dissertation. Along with concerned psychologists, philosophers of science continue to aver the importance of the demarcation problem and characterize pseudoscience in similar ways, for example: (a) resemblance thinking, or confusing superficial similarities and causal relationships (see also Greasley [2010] and the “doctrine of signatures” in “magical medicine” described in Hand [1985]); (b) overall resistance to theory evaluation vis-à-vis rival theories and selective sensitivity to hypothesis confirmation versus disconfirmation (i.e., outright neglect of the scientific method); and (c) consistent lack of theoretical

and evidential progress over time (i.e., stagnation of ideas) relative to more successful scientific research programs (Thagard, 1978, 1980, 1993). The more features that can be identified within particular psychological theories, assessments, and interventions, the more discerning and skeptically cautious mental health professionals and consumers should become (Bunge, 1984, fall; Lilienfeld et al., 2003).

Although developing checklists of scientific and pseudoscientific features in psychology (e.g., see Lilienfeld, 2005) advances the state of inquiry (versus ignoring the problem altogether), overly liberal generation of categorical scientific imperatives to counter the underpinnings of pseudoscience is probably unwise. For example, despite some superficial compatibility of pseudoscientific theory immunization and the principle of methodological tenacity discussed by the controversial “anarchic” philosopher Paul Feyerabend (2002), there may be good reasons to maintain a sound theoretical model despite negative data or what may initially seem to be disconfirming evidence. Specifically, this may be a justifiable course of action in the presence of theory-confirming results from numerous stringent and methodologically sound studies with adequate controls and sufficient statistical power despite a small number of competing, disconfirming results from studies of poorer methodological quality (e.g., disconfirming results from the underpowered study conducted by Quinlan & McCaul [2000] do not single-handedly refute the otherwise well-supported tenets of the transtheoretical model), or even a small number of anomalous findings from well-executed studies. Under such circumstances, some degree of methodological tenacity or “light” immunization may be defensible. However, the principle of tenacity would be rendered untenable in circumstances involving the reversal of this hypothetical

scenario (i.e., few positive results from poorly conducted studies plus many null results from properly designed studies) and would take on an inappropriate form of immunization commonly observed in pseudoscientific “degenerative” research programs (Lakatos, 1970). Thus, hard-and-fast rules are ill advised in this context.

Lilienfeld (1998, fall) has likened science and pseudoscience to Roschian concepts (or “open” concepts; Rosch, 1973) given the absence of unambiguous demarcation criteria, although he draws the helpful analogy that distinguishing day from night remains practical despite the absence of a clear-cut line of division between the two. In the context of this metaphor, one may think of some treatments as falling more squarely within the light of the sun (viz., behavioral and cognitive-behavioral strategies) and others as safely confined to nighttime (viz., recovered memory techniques and TFT). However, some treatment packages (e.g., EMDR) present a more complex picture and appear to inhabit a “twilight zone” of sorts given a combination of both helpful and unhelpful components (Antony & Barlow, 2002; Davidson & Parker, 2001; Lohr, Hooke, Gist, & Tolin, 2003). Of note, an in-depth evidentiary overview of EMDR will be provided at the end of this section to illustrate the nuances of disentangling efficacious and effective therapy ingredients from inert ones under the umbrella of a single intervention.

Of interest, most contemporary definitions of pseudoscience in psychology (e.g., Lilienfeld, 1998, fall; Lilienfeld et al., 2003; Ruscio, 2005; Thagard, 1993) have included both content features (e.g., bizarre claims divorced from evidence) and personal reactions to critics (e.g., burden of proof reversal). However, some philosophers of science (e.g., Derksen, 1993; see also Gardner, 1957) and

psychologists (e.g., Tolin, 2013, May 28) have contended that the primary concern of genuine scientists should be the “pseudoscientists” themselves, because (as noted by Derksen, 1993), “...it is a person, and not a theory or field, who can have scientific pretensions, and who can be blamed for not making good these pretensions” (p. 21). The present author contends that it is likely not entirely possible to parse pseudoscientific content from the actions of pseudoscientists as the two most often go hand-in-hand. After all, it is the person who develops and propounds the content, and it is the person who deploys any number of ill-supported (or, in some cases, intellectually dishonest) escape mechanisms (e.g., ad hoc immunization attempts and ad hominem attacks) when repeatedly confronted with disconfirmatory evidence.

At the same time, however, it is difficult to deny that certain content in itself (generated by a person, of course) should never be characterized as far-fetched within the context of the provisional corpus of scientific knowledge. For example, the foundational content of acupuncture, which entails the insertion of needles at “acupoints” along twelve undetectable “meridians” (supposedly connected to specific human organs and analogically corresponding to the 12 great rivers of China) to stimulate the flow of invisible *qi* (i.e., spiritual “energy” undetectable to physicists) to alleviate illnesses, is utterly denuded of evidence (Bausell, 2007, pp. 113–126; Ernst, 2008; Greasley, 2010; Derry, Derry McQuay, & Moore, 2006; Marcus & McCullough, 2009; O’Connell, Wand, & Goldacre, 2009; Slack, 2010, June) irrespective of the thoughts and behaviors of its most avid historic proponents (e.g., Unschuld, 2003). For this reason, the present author respectfully disagrees with an overly exclusive focus on “pseudoscientists” alone (e.g., Derksen, 1993; Tolin, 2013,

May 28). Rather, the author instead suggests that the interplay of all-too-human cognitive shortcomings, powerful emotional convictions, and dubiously formulated content divorced from the current state of the scientific research (and, in some cases, deliberate charlatanry) interact to foment and perpetuate pseudoscience. In addition, it is important to note that even highly respected, mainstream research scientists are not immune from the same cognitive and emotional biases, and thus likewise may utilize questionable tactics to defend certain hypotheses or theories of professional interest when feeling intellectually threatened, caught off guard, and/or bereft of a polished rejoinder (cf. Derksen, 1993, for a more strongly polarized version of this argument, viz., “It should be stressed that the excessively pretentious and uncritical scientist is not ‘better’ than the pseudo-scientist: he is just more lucky because his theory stands in a critical tradition...” [p. 37]).

The contentious demarcation problem. Much passionate philosophical debate has surrounded the history of distinguishing meaningful and meaningless content (Carnap, 2003), which eventually dovetailed with the perceived meaning and utility of distinguishing science from pseudoscience (Gardner, 1957; Pigliucci & Boudry, 2013; Popper, 2002). In the philosophy of science literature, Larry Laudan (1983) famously relegated the perennial demarcation problem to irrelevancy, dubbing it a “pseudo-problem” and referring to the term pseudoscience as merely a “hollow phrase” doing “only emotive work for us” (p. 125). Instead, he emphasized the central importance of theory confirmation (see also Derksen, 1993; Laudan, 1996). Contemporary objections in clinical psychology include that of Richard McNally (2003), who reviewed a now classic text relevant to the demarcation problem in applied clinical

practice (viz., Lilienfeld et al., 2003). Following in Laudan's (1983) footsteps, McNally (2003) argued that pseudoscience is merely an "inflammatory buzzword" serving no useful purpose for disentangling legitimate from illegitimate scientific endeavors. Instead, the primary concern of psychological researchers should be inquiring about the state of the evidence *for* particular claims to the exclusion of demarcation questions (McNally, 2003), thus arguably emphasizing an astringent form of black box evidentialism (see Shackel, 2013, pp. 421–422) or possibly rote "nose-counting" exercise (Meehl, 1990) with regard to tallying studies with positive outcomes.

Conceding the limitations of a priori plausibility, such as bias introduced by historically predominant yet potentially mistaken scholarly convictions (Ernst, 2003) and the possibility of lapsing into overly dismissive and closed-minded skepticism (Sagan, 1995, January/February), it is difficult to deny that a priori plausibility retains some value for demarcation purposes (see also Beyerstein, 1997, p. 29). For example, when funding agencies are deciding which research to support financially, how would they go about making decisions about treatment outcome studies examining novel interventions (i.e., with no current evidence base) vis-à-vis well-established ones? Is an epistemic free-for-all an economically and pragmatically viable approach whenever a newly proposed intervention emerges? Or would an informed attempt to identify faulty conceptual rationales and mechanisms at the outset prove beneficial for saving valuable time and resources (e.g., those congruent with what is already known to be false and non-scientific, e.g., the principle of analogic correspondences in astrology and herbal remedies; see Greasley, 2010)? Here, an exclusive academic reliance on

strict evidential warrant falls short.

It is the author's contention that there is value in continuing to discuss and clarify the demarcation problem in applied clinical practice given the prevalence of interventions known to be outright ineffective and/or potentially harmful to mental health clients (Lilienfeld et al., 2003). Otherwise, we may waste time and resources unknowingly chasing pseudoscience, all the while lacking an a priori toolkit for separating sense from nonsense. As acknowledged in the philosophy of science literature, there appear to be *two* demarcation problems, namely, a philosophical conundrum and a practical challenge, the latter encompassing the substantial influence on public policy decisions in education, medicine, law, and scientific research funding (Resnik, 2000). As articulated by Resnik (2003) at the conclusion of his philosophical analysis of the demarcation problem, "Our reaction should be that one can distinguish between scientific and unscientific activities even though one cannot rely on a set of necessary and sufficient conditions gleaned from an abstract theory of science to perform this task" (p. 258). In other words, despite inherent logical and philosophical difficulties in delineating unequivocal boundaries between science and pseudoscience, a set of helpful (albeit admittedly limited) criteria (e.g., Bunge, 1984, fall) may be applied on a case-by-case basis, although questions about the supposed practical effectiveness of demarcation criteria would ultimately require meta-scientific study to be addressed appropriately (cf. Faust & Meehl, 2002). Attention will now be turned to an extended review of EMDR, and it is hoped that this discussion will assist in illustrating the complexities of distinguishing science from pseudoscience in clinical psychology.

The Twilight Zone of EMDR: Between Shadow and Substance

Since its inception, Eye Movement Desensitization and Reprocessing (EMDR; Shapiro, 1989a, 1995; Shapiro & Forrest, 2004) has remained a hotly debated psychotherapeutic intervention. Although typically lauded as a speedy remedy for PTSD symptoms (Shapiro, 1989a), EMDR is touted as a “breakthrough therapy” applicable to a wide variety of distressing psychological symptoms according to its founder, psychologist Francine Shapiro (Shapiro & Forrest, 2004). Shapiro (1989a; 1989b; 1994b; 1995) has claimed that EMDR can permanently alleviate the symptoms of PTSD in only a couple of sessions and is more effective than extant cognitive behavioral interventions, albeit in the absence of systematic research evidence. Given what is known about the intransigent and often debilitating nature of severe PTSD (among other anxiety disorders), the glowing testimonials and putatively supportive research associated with EMDR should be met with skepticism, and closer scrutiny of the efficacy and effectiveness of this intervention is warranted.

What is EMDR? The advent of EMDR is not linked to any particular theoretical rationale or compelling logical synthesis, but rather to an anecdotal personal event recounted by Dr. Francine Shapiro. While taking a walk in a park one day in the spring of 1987 and feeling overburdened by distressing thoughts, Shapiro remarked that she instantly felt better after her eyes spontaneously flitted back and forth, thus attributing her improved mood to lateral eye movements (Shapiro & Forrest, 2004). Afterward, she recounted practicing this technique on her friends and acquaintances, many of whom allegedly felt instantly relieved from feelings of anxiety or sadness (Shapiro & Forrest, 2004). Eventually, her experience was translated into

the rhythmic back-and-forth visual tracking technique (*bilateral sensory stimulation*, or BSS) that now defines EMDR (Shapiro, 1999).

The common sequence of steps comprising an EMDR therapy session (cf. Shapiro, 1991; Shapiro & Forrest, 2004) can be summarized as follows. First, the therapist asks the patient to close his or her eyes and imagine the distressing traumatic memory (or a representational image) in vivid detail, much like a typical imaginal exposure. While holding the recalled event in mind, the therapist asks the patient to verbalize any aversive emotional and/or physiological reactions to the event in a sentence. Using a Subjective Units of Distress (or SUDs) scale ranging from 0 (no anxiety) to 10 (extreme anxiety), the patient rates the intensity of psychological distress experienced while imagining the event.

Next, according to the author (cf. Shapiro, 1991; Shapiro & Forrest, 2004), the patient is asked to generate an optimistic statement about the event and subsequently gauge the degree of belief in the positive appraisal using a Validity of Cognition (VoC) scale ranging from 0 (no belief) to 8 (absolute belief). This positive reframing of the traumatic event in conjunction with bolstering its believability constitutes the reprocessing phase of EMDR. While still engaging in the imaginal exposure of the feared scenario, the therapist continues the desensitization process by initiating the technique of BSS, which requires the patient to visually follow the therapist's finger as it sweeps in a lateral, back-and-forth motion approximately 12–14 inches away from the patient's eyes. EMDR therapists typically sweep their finger at a rate of two repetitions per second and total 12–24 repetitions for an average set of repetitions. Exposure to periodic tones in different ears or finger taps can substitute for finger

sweeps if patients are blind or suffer from vision problems (Shapiro, 1994a; 1995). Finally, therapists ask their patients to “blank out” or forget about the traumatic image, breathe deeply, and provide follow-up SUDs and VoC scores after each set of finger sweeps. Finger sweep sets are typically repeated until reported SUDs decrease (e.g., SUDs threshold ≤ 2) and VoC scores increase (e.g., VoC threshold ≥ 6) (see description provided by Lilienfeld, 2008; Shapiro & Forrest, 2004).

The primary difficulty with Shapiro’s claims is not necessarily that her proposed bilateral sensory procedure was borne out of a private experience in a park. Indeed, some ideas emerging from memorable personal experiences, wild hunches, dreams, and other methods of creative or serendipitous freethinking may turn out to be correct upon further testing and corroboration. In his personal correspondence with psychologist Donald Peterson, Paul Meehl pointed out that the German chemist Friedrich Kekulé’s reverie of a snake consuming its own tail (viz., an *ouroboris*) assisted with solidifying the Lewis structure for benzene that remains accepted by chemists to the present day, although other non-dream-related evidence came to bear on this hypothesized structure well before the daydream (Peterson, 2005, pp. 67–68). Unfortunately, however, it is not possible to know with precision from the history of science how frequently such creative hunches are associated with accurate versus inaccurate scientific findings. That is, we are most likely primarily aware of the successful “hits” as opposed to the “misses,” with instances when scientists’ creative ideas turned out to be wrong relegated to the dustbins of history. Thus, the dubiousness of the main tenets of EMDR does not necessarily lie in how they were generated per se. As noted by Thagard (1978, p. 225), “Origins are irrelevant to

scientific status,” although probably not *totally* irrelevant in this author’s opinion (see previous comments on a priori plausibility concerns, plus this is a logician’s argument that, while logically correct, merits testing for its accuracy from the standpoint of empiricism). Rather, the central problem is that the proposed mechanisms of action (and their hypothesized effects) have repeatedly failed to withstand the rigors of scientific testing.

How does EMDR purportedly work? There is a clear consensus among most critics of EMDR that the underpinning theoretical rationale is poorly elucidated and does not square with what is known about the etiology, maintenance, and alleviation of pathological trauma and anxiety (e.g., Keane, 1998; Lilienfeld, 2008; Lohr et al., 2003). As Keane and Barlow (2002) noted, although these criticisms are not grounds for total a priori dismissal of the potential utility of EMDR, it is important to consider the utility of sound conceptual foundations supported by previous research, which in turn may assist with formulating plausible hypotheses and predictions.

To her credit, Shapiro (e.g., 1994b; 1995) has attempted to clarify how EMDR works, although its fit with current understanding of neuroscience and models of cognitive behavioral change is highly questionable (Keane & Barlow, 2002). Specifically, Shapiro (1995; cf. Shapiro & Forrest, 2004) posited that the mechanism behind EMDR relies on accelerated information processing (AIP), an explanatory model purportedly based on neuropsychological principles. In brief, a given traumatic event is thought to impinge upon the nervous system in such a way that distressing information associated with the event becomes encoded without being processed, resulting in neurobiological “blockages.” Traumatic memories are thus improperly

stored and must undergo more adaptive reprocessing and assimilation in the brain through EMDR techniques, such as the back-and-forth eye movements. BSS purportedly expedites the neuropsychological processing of traumatic material by moving information more efficiently through memory networks, much like unclogging a clogged pipeline. Through this dynamically activated processing system, the traumatic content is unlocked. Of note, this explanation can be characterized as a reification fallacy (cf. Gabel, 1976), or concretizing a conceptual metaphor as a psychophysical mechanism of action.

Shapiro (1994b) has also conjectured that beneficial effects of EMDR may result from mimicking eye movements similar to those observed during rapid eye movement (REM) sleep, which supposedly aid in the “processing” of traumatic memories previously inaccessible to the “conscious mind.” However, no proposed neuropsychological mechanism driving the proposed “processing” or how this could alleviate distress has been proposed, and there is of yet no research demonstrating that brain activity associated with undergoing EMDR reflects brain activity during REM sleep (Lilienfeld, 2008). Also of note is the disconnect between involuntary REM sleep eye movements and the smooth, voluntary visual tracking of stimuli in EMDR (Lohr, Tolin, & Lilienfeld, 1998), a qualitative biological comparison gap never broached by EMDR advocates. In addition, Shapiro’s contention that traumatic memories can be repressed or blocked (i.e., in a manner different from forgetting) is itself a deeply controversial claim that lacks a foundation of systematic scientific research support outside of clinical folklore and confected anecdotes (McNally, 2004).

Loose boundary conditions and rapid dissemination. As noted earlier,

EMDR was purportedly crafted for the primary purpose of alleviating the symptoms of PTSD (e.g., Posmontier, Dovydaitis, & Lipman, 2010), although its boundary conditions remain loosely delineated. EMDR has also been recommended for the treatment of various specific phobias (de Jongh, Broeke, & Renssen, 1999), including claustrophobia (Lohr, Tolin, & Kleinknecht, 1996) and spider phobia (Muris, Merkelbach, van Haaften, & Mayer, 1997), as well as chronic pain (Grant & Threlfo, 2002), eating disorders (Hudson, Chase, & Pope, 1998), intellectual disabilities (Rodenburg, Benjamin, Meijer, & Jongeneel, 2009), generalized anxiety disorder (Gauvreau & Bouchard, 2008), deviant sexual arousal (Ricci, Clayton, & Shapiro, 2006), “morbid jealousy” (Keenan & Farrell, 2000), various addictions (Marich, 2010), and “excessive grief, rage, and guilt” (Shapiro, 1991, p. 135). A paucity of research findings demonstrating efficacy for specific disorders and emotional difficulties has not restrained some psychologists from making bold assertions, such as the claim that EMDR is effective in treating ADHD (Tinker & Wilson, 1999) and can improve emotional intimacy in couples’ therapy in the form of Eye Movement Relationship Enhancement (EMRE; Protinsky, Flemke, & Sparks, 2001).

The vociferous positive claims of EMDR proponents about treating PTSD appear to drown out more sober appraisals of existing data. For example, Shapiro (1989b; Shapiro & Forrest, 2004) declared that a single 50-minute EMDR session could substantially reduce or even permanently alleviate PTSD symptoms, a strong claim that has not been substantiated in the literature described next in this review. EMDR was even featured during a 1995 ABC News *20/20* segment (Walters, 1995), which relied on appeals to authority (i.e., the proclaimed expertise of psychologist

Stephen Silver) in lieu of research findings to portray EMDR as a radical breakthrough treatment for various psychological disorders. Bessel van der Kolk, a renowned Boston University professor of psychiatry, reported to the *Los Angeles Times* in 2002: “EMDR sounds like utter nonsense, but this weird thing has a profound effect on people” (Marsa, 2002, March 25). Other dramatic and colorful claims include the supposed ability of EMDR to “pinpoint a specific trauma and target that like a laser beam” (Marsa, 2002, March 25).

By the mid-1990s, over 14,000 psychotherapists had been officially trained to administer EMDR in the United States and abroad (Bower, 1995). Although exact dissemination statistics are uncertain, most likely due to the secretiveness surrounding EMDR training in general (Koocher & Keith-Spiegel, 2008), Shapiro (2004) claimed that approximately 20,000 mental health professionals had been trained in EMDR as of 2004, suggesting a possible increase in dissemination about a decade after the Bower (1995) survey. The popularity of EMDR is unlikely to wane anytime soon, especially in view of the World Health Organization’s (2013; see Recommendation 14) recent public recommendation that CBT, EMDR, and stress management should be considered for children and adults diagnosed with PTSD. Amidst the prematurely aggressive dissemination of EMDR and media hype surrounding its purportedly remarkable healing qualities, critical questions are lost in the fracas: Does EMDR *actually work* as well as is commonly claimed? And if so, does it work in accordance with its *proposed mechanisms of change*?

What is the state of the evidence? Shapiro (1989a) conducted the first study examining the effects of EMDR using a control group. In this investigation, 22

participants, all having experienced a potentially traumatic event, were randomly assigned to an EMDR condition or an imaginal exposure condition without BSS. As hypothesized, the EMDR group evidenced significant reductions in SUDs scores and elevations in VoC scores compared to the imaginal exposure group. Although superficially plausible, these results and other similar findings have been roundly criticized (cf. Acierno et al., 1994; Lilienfeld, 2008) on the grounds that a double-blind procedure was lacking (e.g., Shapiro conducted both treatments herself and personally elicited patient ratings), thus possibly resulting in demand characteristics and/or a Pygmalion effect; exposure times differed across treatment groups; and criteria for terminating distressing imaginal exposures differed across treatment groups (viz., only the EMDR patients were allowed to stop the exposures contingent upon SUDs reductions).

Another major methodological weakness characterizing Shapiro's earlier studies (e.g., Shapiro, 1989a) is the absence of control groups (see Lilienfeld, 2008). Lacking a control group results in a failure to account for the hypothetical counterfactual, or how the patient's symptoms and distress would have fared in the absence of treatment administration (Dawes, 1994). For reasons often articulated in methodology and design texts (e.g., Campbell & Stanley, 1963), drawing strong conclusions about unique treatment effects from uncontrolled studies is problematic (e.g., inability to draw sound causal inferences, difficulty ruling out placebo effects, history effects, maturation, instrumentation, regression to the mean, spontaneous remission, etc.). In addition, combinatory treatments with impure independent variables (e.g., EMDR plus relaxation plus exposure) obfuscate unique effects

attributable to BSS, which remains a key component of EMDR purported to substantially decrease psychological distress (Shapiro & Forrest, 2004).

Lohr and colleagues (1998) reviewed 17 group-design investigations of sounder methodological quality and rigor (e.g., inclusion of random assignment and dismantling designs) compared to previous uncontrolled designs (e.g., Shapiro, 1989b). These authors uncovered systematic evidentiary trends directly contradicting claims of superior efficacy made by EMDR advocates, such as (a) effect size equivalence (Cohen's $d = .90$) across EMDR and non-EMDR exposure treatments; (b) lack of control for therapist by treatment confounds (e.g., therapist enthusiasm and allegiance to EMDR); (c) overreliance on participants' verbal reports of feeling better in the absence of behavioral and physiological measures; (d) lack of significant differences between EMDR and exposure controls (with stationary eye analogue) on behavioral or physiological indicators when they were used (e.g., heart rate, skin conductance, and blood pressure); and (e) lack of significant differences in reported symptom (e.g., Mississippi PTSD Scale) and associated distress (e.g., SUDs) reduction rates across EMDR versus exposure control conditions over time (post-treatment to six-month follow-up).

In the PTSD studies that Lohr and colleagues (1998) examined, although exposure controls and EMDR both yielded better outcomes on SUDs ratings, heart rate, and PTSD symptom ratings compared to no-exposure control groups, EMDR and exposure controls did not significantly differ on any outcome measures. In the reviewed panic studies, EMDR was more efficacious than no treatment but equivalent to no-movement bilateral stimulation analogues. In the reviewed specific phobia

studies, although self-report fear reductions were significantly greater in EMDR conditions in some comparisons (i.e., Muris et al., 1997), participants assigned to in vivo exposure conditions showed greater improvement on avoidance indicators than participants in EMDR conditions, even when EMDR therapists' clinical experience surpassed that of the in vivo therapists.

The majority of Lohr and colleagues' (1998) conclusions from their comprehensive review have been independently corroborated in earlier (e.g., Foa & Meadows, 1997) as well as later (Albright & Thyer, 2010) studies and reviews. In the context of phobic avoidance, Antony and Barlow (2002) reported little to no behavioral (e.g., avoidance ratings made by researchers) or physiological (e.g., lower heart rate and blood pressure) evidence to corroborate patients' verbal reports of fear reduction. In other words, objective indicators have not been utilized to bolster the beneficial impact of EMDR on symptom alleviation. Rather, positive effects remain confined primarily to patients merely saying that they feel better, thus failing to meet triangulation standards (cf. Campbell, 1956).

In a recent study of 74 female rape victims with chronic PTSD symptoms (Rothbaum, Astin, & Marsteller, 2005), there were no significant differences in PTSD symptom reductions between EMDR and prolonged imaginal exposure (PE). This study had the added benefit of including self-report instruments (e.g., the PTSD Symptom Scale-Self Report and Impact of Event Scale-Revised) as well as structured and semi-structured clinical assessment interviews (e.g., the Clinician-Administered PTSD Scale and SCID) conducted by independent raters blind to treatment condition. Treatment integrity ratings were also included.

It is important to note that a number of studies questioning EMDR efficacy and effectiveness can be classified as experimental dismantling studies⁷ (e.g., Cahill, Carrigan, & Frueh, 1999; Renfrey & Spates, 1994), which are sometimes called additive or subtractive designs (Lohr et al., 2003). These studies entail the removal of specific treatment elements (i.e., BSS) from otherwise intact treatment packages, which can be compared to treatments containing the component in question to determine if removal diminishes treatment efficacy (cf. Hart, Fann, & Novack, 2008). These types of designs thus aid in disentangling specific from non-specific treatment factors. Studies pitting EMDR against exposure control conditions lacking BSS have detected no significant differences in treatment effects as shown by standardized psychological and physiological measures (see Boudewyns, Stwertka, Hyer, Albrecht, & Sperr, 1993). Renfrey and Spates (1994) discovered that when the eyes are fixed on a stationary light, thus subtracting out the eye sweep technique that is the lynchpin to the bilateral sensory stimulation rationale, the treatment results are no different from those of EMDR including the technique. Follow-up studies using similar no-movement analogues with Australian Vietnam veterans suffering from PTSD have replicated this negative finding (Dewilly, Spence, & Rapee, 1998). Taken together, these results cast serious doubt on the claim that the eye movement component is needed at all in EMDR.

Perkins and Rouanzoin (2002) cogently summarize the manifold methodological, statistical, and conceptual shortcomings plaguing EMDR as follows: (a) lack of a valid empirical framework that coherently predicts and explains observed treatment effects; (b) lack of methodologically sound studies; (c) the possibility of the

file drawer problem (i.e., lack of reporting null results combined with a selective sensitivity to publishing positive results) among EMDR researchers with vested interests in positive outcomes; (d) an accumulation of positively biased reports by a homogeneous group of researcher-clinicians with a personal stake in the matter; and (e) suspected inconsistencies in treatment fidelity across studies. Despite this clear confluence of data pointing to the therapeutic impotence of the eye movement technique, EMDR as a whole continues to be heralded as a “breakthrough treatment” (Shapiro & Forrest, 2004) and has been dubbed potentially effective by both the Department of Veterans Affairs and Department of Defense (2004).

Briefly, with regard to diversity issues, few studies have discussed ethno-racial or cultural factors in the context of EMDR treatment, typically going no further than providing demographics tables in results sections. The vast majority of research participants across virtually all studies cited within this review were Caucasian adolescents and adults, with African Americans comprising a small percentage of participants (e.g., 2-3%). Among the few articles recruiting samples from non-North American populations was an investigation of EMDR involving a small group of Iranian children aged 12-13 (Jaberghaderi, Greenwald, Rubin, Zand, & Dolatabadi, 2004) and a single-session EMDR group intervention conducted with a sample of children aged 4-17 following the calamitous 2003 flooding in Santa Fe, Argentina (Adúriz, Bluthgen, & Knopfler, 2009).

In the Middle Eastern study (Jaberghaderi et al., 2004), 14 Iranian children with histories of sexual abuse were randomly assigned to either an exposure-based CBT condition ($n = 7$) or an EMDR condition ($n = 7$) to be treated for trauma

symptoms. Comparisons of pre- and post-treatment measures of self- and other-reported PTSD symptoms alongside behavioral ratings made by parents and teachers revealed roughly equal efficacy of CBT and EMDR, but the authors concluded that EMDR was more efficient (e.g., fewer sessions and more rapid intra-session SUDs reductions). However, this latter conclusion is questionable given that the CBT condition required a minimum of 10 sessions with heavy psychoeducational requirements, whereas the EMDR condition categorically lacked these requirements with termination contingent on quickly reaching low SUDs thresholds. Similar favorable results were found for the EMDR intervention in the South American study (Adúriz et al., 2009), which recruited 124 schoolchildren who had been forced to evacuate their homes with their families due to severe flooding. Although SUDs ratings significantly decreased from pre- to post-treatment and PTSD symptoms (e.g., intrusion and avoidance symptoms) and remained significantly lower at 3-month follow-up, there was no CBT group or control group included for comparison. Enhancing the methodological quality of and participant recruitment efforts associated with these sorts of studies is critical given the insufficient inclusion rates of ethnic minorities and non-North American samples in efficacy studies and small sample sizes—shortcomings that have continued to hamper generalizability (i.e., external validity) of results to ethno-culturally diverse populations (Miranda, Nakamura, & Bernal, 2003).

EMDR: A concluding summary of the evidence. The aforementioned studies all converge on the same conclusion: The observed effectiveness of EMDR in the extant literature can be reasonably attributed to its imaginal exposure component

and not BSS (Antony & Barlow, 2002; Davidson & Parker, 2001). Evidence supporting the therapeutic contribution of bilateral sensory stimulation components (i.e., lateral ocular movements and alternating visual field stimulation) is either weak or non-existent (Davidson & Parker, 2001; Lohr et al., 2003). When writing their comprehensive review of state-of-the-art interventions for PTSD, Keane and Barlow (2002) asserted that no existing study to date had demonstrated incremental efficacy of EMDR over and above any existing evidence-based treatment for PTSD (i.e., anxiety management training, cognitive restructuring techniques, and imaginal exposures). Nearly thirteen years later, this observation remains unchallenged by methodologically sound studies. Furthermore, compared to individuals who have undergone EMDR, those who receive traditional CBT for PTSD symptoms have attained greater treatment gains as evidenced by both post-treatment and follow-up assessments (Deville & Spence, 1999). Until a coherent synthesis of data contradicts this body of evidence, there would seem to be no defensible rationale for replacing tried-and-true CBT techniques with EMDR.

Of note, this review illustrates EMDR proponents' striking neglect of rigorous scientific inquiry into purported mechanisms of change through various pseudoscientific maneuvers, namely, (a) dispensing with the proper methodological toolkits associated with efficacy and effectiveness research; (b) repeatedly exaggerating unfounded claims that a specific, unsupported treatment element (in this case, BSS) works; and (c) ignoring extant evidence that the impact of BSS appears to be no greater than non-specific or placebo effects at best. Thus, the BSS component of EMDR may be viewed as pseudoscientific as a function of its inertness in the

inseparable context of intransigent confirmatory attitudes of avid proponents as well as the scientific implausibility of the proposed theoretical underpinnings. This illustrates a possible compatibilist stance between the distracting debates on whether we should focus exclusively on pseudoscientists (e.g., Derksen, 1993; Tolin, 2013, May 28) or pseudoscientific content (e.g., Bunge, 1984, fall; Lilienfeld et al., 2003). In this sense, the pseudoscience label may provide a concise and informative descriptive heuristic, as it is not being used in a cavalier, dismissive, or inflammatory ad hominem manner.

The burden of elucidating and laying the ground for testing proposed neuropsychological mechanisms of action of EMDR lies with the claimants. In lieu of evasive, ad hoc defensive maneuvering, such as Shapiro's revisionist statement suggesting that eye movement is neither a necessary nor sufficient treatment component (see Lohr, Hooke, Gist, & Tolin, 2003), EMDR proponents should try to develop a concrete set of testable hypotheses nested within a clear, coherent rationale drawing from the extant research literature on anxiety and trauma. However, given the data already scrutinized, it may be argued that such a step would thrust us squarely into a fallacy of misplaced rationalism (Sheaffer, 2008). In other words, EMDR enthusiasts may be attempting to speciously explain an inert, non-existent phenomenon from a position of post hoc rationalization, thus further exacerbating the pseudoscientific practice of immunization from falsification (see Bunge, 1984, fall). In this sense, future research directions are not entirely clear.

Pseudoscience: What's the Harm?

Interventions based on pseudoscience may not merely contain inert components that fail to provide benefits (e.g., as observed with BSS in EMDR)—rather,

they may harm clients. Overall, psychotherapy researchers have paid scant attention to baneful consequences of psychological treatments in recent decades, although within the past 5 years, some literature has focused on raising awareness of potentially harmful treatments (or PHTs; Castonguay et al., 2010) and how to detect and address such effects (Dimidjian & Hollon, 2010). Moreover, some psychologists support the creation of an official list of PHTs similar to existing EST lists (Castonguay et al., 2010; Lilienfeld, 2007).

PHTs were first formally defined by Lilienfeld (2007) as interventions that have resulted in adverse psychological and/or physical effects that are both enduring (i.e., not merely a temporary worsening of symptoms) and have been replicated by independent lines of research. By psychological effects, Lilienfeld (2007) refers to symptom exacerbation plus attenuated improvement rate over the course of care. And by physical effects, he refers to either physical harm (e.g., risk of parasitic infections or injury from aggressive mating behavior in dolphin-assisted psychotherapy; Marino & Lilienfeld, 2007; Samuels & Spradlin, 1995) or death (e.g., as has occurred in rebirthing therapy; Josefson, 2001). Among provisionally identified PHTs are several psychological treatments classed as pseudoscientific, namely, critical incident stress debriefing (CISD), rebirthing therapy (RT), and recovered memory therapy (RMT), which will be reviewed briefly in turn.

CISD is a 3–4-hour, single-session group psychotherapy procedure in which clients openly disclose distressing thoughts and feelings in the aftermath of a potentially traumatic event, assumedly to avoid the onset of PTSD symptoms (Lohr, Hook, Gist, & Tolin, 2003). According to the guidelines of CISD, clients must (a)

participate in therapy no later than 24 to 72 hours after the trauma, (b) discourage one another from leaving the therapy group once it has begun, and (c) discuss possible PTSD symptoms that they may face as a result of the traumatic event (Lohr & Fowler, 2002, summer; Lohr et al., 2003). Of note, CIRD has been found to be consistently ineffective at best and possibly harmful at worst (e.g., worsened PTSD symptoms in CIRD groups compared to assessment-only controls) in treating PTSD symptoms across both meta-analyses (e.g., effect size of $d = -.11$; Litz, Gray, Bryant, & Adler, 2002) and RCTs (e.g., Bisson, Jenkins, Alexander, & Bannister, 1997; Mayou, Ehlers, & Hobbs, 2000). Some researchers have hypothesized that observed PTSD symptom exacerbation at follow-up may be partly due to interference with natural symptom remission (Gist & Woodall, 1995). As opposed to using systematic desensitization (with the goal being gradual habituation) in tandem with practicing alternative adaptive responses to alleviate symptoms associated with distressing aspects of a trauma (as is the case in evidence-based CBT approaches; Foa, Zoellner, & Feeny, 2006), CIRD instead encourages therapists to ask direct questions about the worst aspects of the trauma during the reactions/cathartic ventilation phase (i.e., shortly after the trauma) in the absence of teaching coping techniques (Deville, Gist, & Cotton, 2006). This may be especially problematic for subgroups of patients who struggle with dysregulated hyperarousal (Deville et al., 2006).

RT is a type of attachment therapy that has been flagged as potentially dangerous depending on how it is practiced (Mercer, 2008). RT was introduced in 1974 by Leonard Orr, a self-proclaimed pioneer of the New Age movement who continues to lead the Rebirthing Breathwork movement and claims to have unlocked

the secrets of physical immortality (cf. Orr's 1998 book, *Breaking the Death Habit: The Science of Everlasting Life*). Orr proposed that human birth is always accompanied by a fear of suffocation triggered by the premature severing of the umbilical cord, which supposedly damages the person's "breathing mechanism" and embeds panic deep into the subconscious mind (Singer & Lalich, 1996, pp. 42–43). This repressed fear purportedly resurfaces in the form of both psychological (e.g., anxiety, depression, and low self-esteem) and physical ailments (e.g., allergies, weight problems, and cancer), and Orr proclaimed that several 2-hour "rebirthing" sessions (e.g., practicing patterned breathing techniques while floating or snorkeling in a hot tub) is usually sufficient for healing these conditions as well as fostering psychic abilities (Singer & Lalich, 1996, pp. 42–44).

Other variants of rebirthing include dramatic "recapitulations" (as described in the psychoanalytic attachment literature) of the birth process by wrapping up clients, especially young children with developmental problems, in carpets or blankets while squeezing them, taunting them, and encouraging them to struggle free (Lilienfeld, 2007; Mercer, 2008). Not only do these preposterous and needlessly abusive techniques lack any supporting research evidence (e.g., no RCTs have been conducted, and no evidence of efficacy or effectiveness can be found in the peer-reviewed literature), but they also have resulted in reported serious injuries and even deaths, including the asphyxiation of a 10-year-old girl in Colorado in 2001 (Mercer, 2008). Of note, the two social workers responsible for the girl's death were sentenced to 16 years in prison, and a new legal mandate known as Candace's law, which prohibits restraint in psychotherapy, was passed in Colorado and North Carolina shortly

thereafter (Josefson, 2001; Mercer, 2008).

Finally, RMT entails the use of various highly questionable, poorly supported techniques (e.g., hypnosis, “age regression,” sodium pentathol administration, guided imagery, and/or therapist interpretations of symptoms) to recover memories of traumatic past events assumed to have taken place during a client’s childhood (Lynn, Loftus, Lilienfeld, & Lock, 2008). As summarized elsewhere (Loftus, 1993; Lynn, Lock, Loftus, Krackow, & Lilienfeld, 2003; Lynn et al., 2008; Singer & Lalich, 1996; Singer & Nievod, 2003), these memories supposedly become “repressed” deep into the unconscious mind due to the intense trauma and emotional pain associated with aversive early experiences. The supposed recovered memories are frequently of questionable veracity and have a number of bizarre cottage industry therapy movements associated with them (e.g., Satanic ritual abuse therapy, alien abduction therapy, past-life regression, and entities therapy; see Singer & Nievod, 2003). RMT advocates (e.g., Fredrickson, 1992) posit that if these memories are not recovered and emotionally processed in therapy, they may result in a host of baleful life consequences, including chronic psychological distress, tumultuous relationships, professional failure, and in extreme cases, personality fragmentation resulting in dissociative identity disorder (formerly multiple personality disorder). Not surprisingly, evidence supporting the efficacy and effectiveness of RMT is either woefully inadequate or non-existent at the present time, and RMT memory retrieval techniques (esp. hypnosis) in particular have been thoroughly debunked with regard to reconstructing accurate representations of past events (see Lynn et al., 2008, for a review).

Sadly, RMT has resulted in a multitude of wrongful prosecutions of and civil lawsuits against parents who allegedly sexually abused their children (Loftus, 1995; Maran, 2010; Wakefield & Underwager, 1992), evidence for which was gathered during therapy sessions using the questionable procedures mentioned in the previous paragraph (see also Loftus & Ketcham, 1994). Of note, many of these unfortunate accusations occurred in the heat of a mass hysteria known as the “Satanic Panic,” which swept across the United States during the 1980s and early 1990s (Nathan & Snedeker, 2001; Victor, 1993). During this period, popular daytime television talk shows (e.g., Geraldo Rivera, Oprah Winfrey, Sally Jesse Raphael, and Donahue) dramatically belabored the supposed existence of a large, clandestine sect of Satanists gleefully involved in routine macabre activities, including human sacrifices, animal mutilations, desecrations of religious buildings, cannibalism, the distribution of illicit drugs, kidnapping, pedophilia, and the production and distribution of child pornography (Victor, 1993). In professional mental health circles, popular therapy manuals, such as Bass and Davis’ (1990) *The Courage to Heal Workbook*, further perpetuated dangerous and irresponsible claims, such as the classic assertions, “If you are unable to remember any specific instances... but still have a feeling that something abusive happened to you, it probably did” (p. 21), and, “If you think you were abused and your life shows the symptoms, then you were” (p. 22). One particular author (Cautin, 2011) went so far as to attribute a seismic exacerbation of the research-practice gap to the recovered memory controversy of the 1980s and early 1990s, implying that the overall message of the abovementioned lawsuits (i.e., the public portrayal of the sloppy clinician versus the pristine scientist) further injured the

already contentious relationship between clinical researchers and practitioners.

It should be clarified at this point that not all treatments categorized as pseudoscientific are necessarily harmful, and not all harmful treatments have the trappings of pseudoscience per se. For examples of the first point, no direct deleterious psychological or physical effects have yet been documented for the BSS component of EMDR. Nor have any such effects been observed for TFT, which involves lightly tapping meridian points at various locations on the body with one's fingers while voicing positive self-affirmations, the purported goal being to unblock thought field "energy" obstructed by trauma (Feinstein, 2008). As for the second point, for a large subset of people experiencing typical bereavement (e.g., dysphoria following the death of a close family member), grief counseling appears to be associated with a clear deterioration of psychological and behavioral functioning post-treatment (compared to no treatment) according to a meta-analysis of RCTs (Neimeyer, 2000). However, grief counseling itself does not consist of pseudoscientific approaches per se (e.g., fostering social and familial support, reinforcing meaning making associated with death, and reflecting on positive memories) and may yield more beneficial effects for complicated grief (see Allumbaugh & Hoyt, 1999; Altmaier, 2011; Bonanno & Lilienfeld, 2008).

These considerations aside, practicing clinicians serve their clients best by avoiding harm to the extent possible (i.e., when it is foreseeable; APA, 2002) and using treatments supported by scientific evidence. Deferring to personal preferences for treatments (e.g., TFT for PTSD) and tenaciously maintaining that a suboptimal treatment choice is justified because it has not been shown to be associated with harm

arguably constitutes unethical clinical practice, especially when the current state of the evidence is ignored. Unfortunately, however, the current APA Ethics Code (2002) explicitly confers equal status to professional judgment and peer-reviewed research findings in both clinical practice and pedagogical decisions, which is unjustified given the large clinical decision-making literature.

Clinical Decision Making

The clinical method and the actuarial method (Dawes, Faust, & Meehl, 1989) are the two central decision-making methodologies for predicting behavior or outcomes in clinical psychology. The clinical method entails an “in the head” or impressionistic synthesis of information to arrive at a conclusion, whereas the actuarial method relies on empirically-established relations between information or data (e.g., frequencies) and the outcome of interest, which are analyzed formally (e.g., using mathematical formulae or tables; Grove & Meehl, 1996) to reach a conclusion or probability statement. Since the mid-twentieth century, literally hundreds of studies in the decision-making literature have converged on the conclusion that the actuarial method almost always equals or exceeds the clinical method in accuracy (e.g., when predicting the presence or absence of a diagnosis), sometimes by a small margin and sometimes a considerable margin, and hence is the superior method overall (Goldberg, 1965; 1968; 1970; Dawes, 1971; Einhorn, 1972; Meehl, 1954; Sawyer, 1966). This clear-cut trend has continued to emerge in meta-analyses as well (e.g., Aegisdóttir et al., 2006; Grove et al., 2000).

It is evident across this vast literature that many applied clinicians either ignore or inappropriately countervail readily available actuarial data and instead defer to

clinical judgment. To the author's knowledge, there are no formal investigations of individual cognitive and psychosocial factors undergirding this irrational decisional intransigence, although Meyer, Baker, and Baker (2012, March) summarized possible misguided epistemological justifications for overreliance on the clinical method. These include (but are not limited to) the following: (a) the fallacy of commensurate complexity, or the erroneous notion that human behavior is so complex that only equally complex methodologies could provide accurate prediction (cf. Faust, 2007); (b) the fallacy of *argumentum ad experientiam*, or stubbornly deferring to one's own self-assumed clinical expertise and experience as the best method for predicting outcomes; (c) an illusion of perfect predictability or consistent error avoidance via faulty methods, also framed as a steadfast resistance to accepting error to yield less error (see Einhorn, 1986); and (d) avowal of the ecumenical decree that no controversy exists—one can simply integrate both methods (Grove & Meehl, 1996), presumably even if they yield contradictory or mutually incompatible outcomes. However, such postulated reasons for resistance to the use of actuarial methods remain to be tested and elucidated through a systematic program of research with clinical practitioners as the target population.

Despite the robust limitations of expert clinical judgment, including studies illustrating a weak relationship between clinical experience and accuracy⁸ (see Dawes, 1989; Lilienfeld et al., 2003), commensurate levels of accuracy among novices and experts on a number of judgment tasks (especially when clinicians lack sound scientific evidence or ignore it; Goldberg, 1968; Weck, Weigel, Richtberg, & Stangier, 2011), and observed deterioration of knowledge once the practitioner completes

formal education and training (Vollmer, Spada, Caspar, & Burn, 2013), this has not stopped many clinical psychology authors from professing an exaggerated faith in clinical expertise. Some even go so far as to declare by fiat that many years of clinical experience render one a more effective therapist in the absence of supporting data (e.g., Betan & Blinder, 2010). For example, Overholser (2010) proposed the following domains of clinical expertise: (a) possession of a terminal degree in psychology, (b) accrual of many years of direct clinical experience administering treatments and/or assessments, (c) demonstration of strong clinical skills in specific clinical applications, (d) possession of an advanced graduate degree in a relevant specialty area, and (e) national visibility in a specific professional community.

Lamentably, a call for accurate knowledge of the relevant scientific literature (and a fortiori, knowledge of effective interventions and their applications) is nowhere to be found in this formulation of superior clinical acumen. Although the present author disagrees with the insistence that applied practitioners become involved in formal research endeavors (e.g., Marten & Heimberg, 1995), it would be ethically indefensible to maintain that they are justified in neglecting to keep abreast of the current state of the evidence for relevant assessments and interventions routinely used in their clinical practice. This is especially true given that the estimated half-life of knowledge in clinical psychology is approximately 11 years with an anticipated future decline to 9 years (Neimeyer, Taylor, Rozenky, & Cox, 2014). In other words, after roughly 9–11 years pass following the completion of their doctoral degree, 50% of what psychologists learned in graduate school will become obsolete. Thus, ironically, the very encouraging finding that knowledge is advancing rapidly in psychology is

potentially offset by the extent to which such knowledge may be routinely disregarded or replaced with pseudoscientific or insufficiently validated beliefs.

With reference to explicit overvaluation of clinical experience and expertise, Betan and Binder (2010) recently introduced *metabolizing theory*, which presumes that “expert” therapists differ from novices in that they engage in a flexible, adaptive, and accurate intuitive synthesis of theoretical and clinical knowledge outside of conscious awareness when formulating case conceptualizations and treating clients. In contrast, novice or non-expert therapists (i.e., those with far fewer years of accrued clinical experience) presumably lack a metabolic, intuitive grasp of core concepts. Despite their apparently genuine convictions, the authors did not reference a single study supporting these bold assertions. Such assertions are likely better categorized as questionable conjecture or hypothesis as opposed to a theory given the absence of supporting evidence and the overwhelming presence of negative findings that contradict such optimistic pronouncements about the superior levels of accuracy achieved through experienced clinical judgment (Garb, 1998). Strongly worded yet evidentially hollow claims of this sort pertaining to clinical expertise may inadvertently instill a sense of professional complacency and provide a disincentive to keep up with relevant scientific research and/or utilize evidence-based tools.

All of this is not to say, however, that clinical judgment should be the nemesis of applied psychologists. Clinical intuition, expertise, and judgment play valuable roles in the context of discovery (e.g., hypothesis generation) and should not be considered second-class given their role in advancing the field (Lilienfeld, 2010; Chambless, 2014). Practitioners who draw from their rich clinical experiences in

formulating case studies or presentations and subsequently utilize this knowledge to propose clear, sensible, and testable hypotheses for future research are constructively contributing to the true integration of science and practice (Lowman, 2012). Some of these hypotheses turn out to be correct (or nearly so), and at present, it is difficult to conceive of other ways by which novel clinical ideas could be generated. As a case in point, Chambless (2014) recently described how Aaron Beck's seminal cognitive conceptualization of depression began with a series of loose-knit observations of dysphoric patients and their reported thought content, which gradually evolved into a more streamlined and systematized program of research. In this same vein, mutually informative dialogue between practitioners and researchers may help close the widening research-practice gap by having practitioners assist as research "problem finders," whereas researchers would primarily serve as "problem solvers," although these roles blend together for some psychologists (Chambless, 2014; Goldfried et al., 2014). Nevertheless, these same mental processes that shape initial clinical impressions and fledgling research agendas are certainly not immune from the pervasive influence of cognitive biases (e.g., confirmation bias), especially in the context of justification (Lilienfeld, 2010).

Cognitive Characteristics of Clinicians

As acknowledged in the decision-making literature, clinical psychologists routinely face complex and ambiguous scenarios that demand fairly rapid decisions about specific courses of action (e.g., diagnosis, suicide and violence risk assessment, intervention selection, and prognostic forecasting; Olthmanns & Klonsky, 2007). Unfortunately, psychologists in applied practice are rarely given immediate corrective

feedback on their decisions about specific scenarios followed by opportunities for repeated practice (i.e., *deliberate practice*; Ericsson, 2005; Lewandowsky, Little, & Kalish, 2007), which are known to be invaluable conditions for establishing expert levels of decisional accuracy (Ashby & O'Brien, 2005). An example of a field in which perceptual expertise is developed through repeated deliberate practice is ornithology, where experts gradually learn how to quickly and accurately identify birds at subordinate levels of representation (Krigolson, Pierce, Holroyd, & Tanaka, 2008). A skilled ornithologist, for example, would be able to identify an American flamingo as belonging to the species *Phoenicopterus ruber* as rapidly as a novice could identify it as a pink flamingo. In contrast, clinical practitioners typically cannot quickly and accurately diagnose a client based on readily observable pathognomonic features, not necessarily due to lack of competence, but also attributable to the complex heterogeneity of mental illness manifestation⁹ (Seaton et al., 1999). Furthermore, they often lack access to accurate corrective diagnostic feedback in their work settings in the same way that an ornithologist could compare observed birds to a textbook of representative exemplars, although professional consultation may assist with diagnostic accuracy depending on the nature and quality of the feedback. The feedback they do receive often contains a fairly large error component, which can greatly diminish the benefits of experience and easily foster mistaken belief (Dawes, 1989).

Faced with these environmental pressures (e.g., obstacles to proper experiential learning in a fast-paced work environment), clinical psychologists are understandably susceptible to a pernicious host of cognitive biases and heuristics, which may result in

suboptimal clinical decision-making procedures, overconfidence, an illusion of learning, and decreased judgmental accuracy (Arkes, 1981; Dawes, 1994; Dawes et al., 1989; see also Tversky & Kahneman, 1974). As Meehl (1993) more forcefully stated, “It is absurd, as well as arrogant, to pretend that acquiring a Ph.D. somehow immunizes me from the errors of sampling, perception, recording, retention, retrieval, and inference to which the human mind is subject” (pp. 728). Generally speaking, cognitive heuristics can facilitate rapid adaptive choices in an overwhelmingly complex world, but they may also muddle perceptions and reinforce false beliefs (Kahneman, 2011; Tversky & Kahneman, 1974).

Cognitive and personologic variables contributing to clinician susceptibility to pseudoscientific beliefs are not well researched or formally understood, but it is plausible to hypothesize that they do so largely because of the ubiquitous human vulnerability to cognitive biases (and, more generally, irrationality) alluded to in previous paragraphs¹⁰. For example, following their initial attraction to and utilization of pseudoscientific treatments due to whatever combination of biases and proclivities, clinicians may subsequently fall into what social psychologist Anthony Pratkanis (1995, July/August, p. 21) calls a *rationalization trap*. This entails developing a gradual personal commitment to (or, stated another way, building an emotional investment in) the core principles of the intervention, which may be facilitated by various cognitive biases and stressors (e.g., confirmation bias and cognitive dissonance). Of course, cognitive bias susceptibility and illogical thinking may be more acutely amplified in some individuals for whatever reasons. Some studies have found, for example, that individuals who believe more strongly in the paranormal tend

to commit more logical and probabilistic judgment errors than their more skeptical peers (see Majima, 2012).

Interestingly, in tandem with the rationalization trap, some clinicians may become so committed to particular theoretical frameworks that they begin to experience the following phenomenology: (a) ownership of their position, (b) perceiving their position as an extended part of their self-concept, and (c) subsequently perceiving any criticisms of the underlying conceptual and/or evidentiary apparatus as an attack on *themselves* (e.g., akin to an erroneously perceived *ad hominem* attack; see de Dreu & van Knippenberg [2005] for preliminary experimental evidence for this possibility). Involvement with “granfalloon” (Vonnegut, 1973, as cited in Pratkanis, 1995, July/August, p. 22), or “proud and meaningless” in-groups emphasizing a cohesive social identity associated with shared beliefs and jargon, may further reinforce wayward clinicians’ commitment to favored interventions and/or theoretical frameworks (e.g., see McNally [1999] for a telling description of the EMDR Institute, Inc., and its members’ activities). This may also serve to isolate them from informed skeptics by reinforcing perceptions of them as hostile out-group members (Pratkanis, 1995, July/August).

Regrettably, exposure to general higher education alone does not appear to be a sufficient buffer against the proliferation of pseudoscientific and paranormal beliefs. For example, a paranormal beliefs survey given to 133 allied health students (e.g., undergraduate and graduate students in the fields of physical therapy, medical technology, and health administration) from two universities (including an Ivy League university) revealed the following: belief in extrasensory perception (46%), perceived

legitimacy of chiropractic medicine (36%), and claims of telepathic experiences (25%) (Duncan, Donnelly, Nicholson, & Hees, 1992). However, educational interventions targeting domain-specific critical thinking skills do hold promise for mitigating at least some influences of faulty reasoning associated with questionable beliefs. Relative to a general psychology undergraduate course, a cooperative learning course specially designed to assist students with critically scrutinizing paranormal claims resulted in statistically meaningful gains in abilities to detect and analyze reasoning flaws ($d = 1.2$) as well as generate rival alternative scientific explanations for phenomena of interest ($d = .84$) (Wesp & Montgomery, 1998). Additionally, significant reductions in paranormal beliefs found among students in an undergraduate science seminar and a pseudoscience seminar (compared to a quasi-control group) were maintained over a 2-year period following the end of the courses (Dougherty, 2004; Morier & Keeports, 1994). Perhaps similar improvements might be attained were clinical psychologists exposed to educational interventions designed to strengthen relevant critical thinking skills and the ability to differentiate scientific from pseudoscientific therapeutic claims, although this has yet to be formally tested.

Baker and colleagues (2008) argued that the following characteristics of practicing clinical psychologists contribute to their professional ostracization (i.e., being “crowded out” of mental health service delivery by primary care physicians and Masters-level clinicians) and overall weak impact on modern healthcare: (a) overvaluation of clinical experience at the expense of weighing available scientific evidence, (b) persistent use of poorly supported interventions in place of well-supported ones, (c) lack of adequate education and training in scientific foundations,

and (d) ambivalent feelings about science and its utility in applied clinical psychology. Of note, factors “a” (intuitive decision-making preferences) and “d” (ambivalence or, in some cases, hostility toward science) in particular have been repeated foci of discussion in the pseudoscience literature in clinical psychology (Lilienfeld, Lynn, & Lohr, 2003) and speak to important cognitive-emotional preferences of applied psychologists. However, as Garb and Boyle (2003, p. 30) noted, no studies to their knowledge (or the current author’s knowledge) have attempted to qualitatively or quantitatively investigate these features, especially as they relate to specific behavioral consequences (e.g., the aforementioned factors “b” and “c”) pertinent to professionally and ethically responsible mental health care.

In addition, apart from a small number of recent surveys, which are limited in scope (e.g., Sharp et al., 2008), there is limited understanding of how critical thinking skills and cognitive styles relevant to clinical decision-making relate to knowledge about evidence-based and pseudoscientific treatments. These variables are worthy of further research given the proliferation of highly questionable and, in many cases, suboptimal interventions offered to mental health clients. As Lilienfeld (2010, p. 283) noted in his review of the treatment literature, which is only a small snapshot of the pervasiveness of the problem at hand, (a) tens of thousands of clinicians have received EMDR training (cf. Bower, 1995; Shapiro, 2004); (b) most clients suffering from anxiety, mood, eating, and autism spectrum disorders do not receive scientifically supported psychotherapies, and (c) increasing numbers of clients suffering from mental illness receive unsupported and questionable interventions, such as “energy therapies” (e.g., TFT). Although a number of studies lend credence to the claim that

poor critical and scientific thinking skills are associated with stronger beliefs in pseudoscience and superstition among high school and college-age students (e.g., Bennett, 1991; McKenzie, 1986), this has yet to be examined among applied psychologists in active clinical practice.

Study Aims and Hypotheses

Given the absence of an existing measure, the primary objective of this study was to develop a questionnaire (viz., the Clinical Attitudes and Knowledge Questionnaire, or CAKQ; see “Part B” of the Appendix) intended to appraise specific clinical knowledge domains and attitudes toward science among licensed, doctoral-level practitioners of clinical psychology. This aim was pursued through generating an initial item pool designed to detect the presence of knowledge pertaining to (a) legitimate as well as questionable treatment techniques used in contemporary clinical practice (e.g., evidence-based versus poorly supported or, from some perspectives, pseudoscientific approaches [see Lilienfeld, Lynn, & Lohr, 2003]); (b) relevant clinical psychology research areas (e.g., current controversies that impact applied practice); and (c) clinical judgment and decision-making procedures. A preliminary scale consisting of items addressing practitioner attitudes toward science in clinical psychology was also created. A secondary study aim was to ascertain whether psychologists’ professed knowledge varied in relation to years involved in clinical practice. Preliminary questionnaire data were collected from practicing clinical psychologists in various professional settings and analyzed to examine relationships among clinical knowledge and relevant cognitive variables.

As this dissertation is the first study of this new questionnaire, direct research

evidence on its latent structure was lacking. Hence, hypotheses were generated based on general themes of item parcels. Excluding the attitudes toward science items from statistical analyses, it was hypothesized (**Hypothesis 1**) that a four-component (or, alternatively, factor) solution would best summarize the data (i.e., knowledge of pseudoscientific interventions, evidence-based interventions, general clinical research, and clinical judgment/decision making).

Regarding the predicted nature of relationships among clinical knowledge and relevant cognitive and information processing variables (i.e., intuitive processing styles and critical thinking skills), it was hypothesized (**Hypotheses 2 and 3, respectively**) that higher scores on the CAKQ (i.e., greater accuracy in answering questions relevant to the research areas listed above) would be significantly associated with (a) higher scores on inference, deduction, and interpretation tasks, and (b) a lower reliance on intuitive thinking styles. Finally, it was hypothesized (**Hypothesis 4**) that clinical knowledge would not significantly vary as a function of clinical experience across participants (i.e., more years of clinical experience would not be associated with higher CAKQ scores).

CHAPTER 3

METHODOLOGY

Participants

Two thousand randomly selected licensed psychologists currently engaged in clinical practice in New England were invited to participate in the study. Inclusion criteria included: (a) current licensure to practice psychotherapy in a New England state; (b) possession of a doctorate in the field of psychology (e.g., Ph.D. or Psy.D.), education (e.g., Ed.D.), science (e.g., Sc.D.), or other related field; (c) current active involvement in clinical practice (e.g., not retired or exclusively involved in research); and (d) the ability to read, write, and understand English.

Of the 2,000 questionnaires mailed, 345 surveys were returned completed in full or part (see further below for details). In addition, 126 unopened survey packets were received by return mail, indicating outdated practitioner addresses (e.g., practitioners who had relocated or were deceased as noted on the envelopes), and 12 psychologists contacted the student investigator directly by telephone or email to decline participation for various reasons (e.g., they were retired, no longer in applied clinical practice, no longer licensed, or did not have time to complete the survey).

Among the 345 individuals who returned surveys, 335 were currently practicing licensed psychologists, and 10 were retired psychologists no longer in applied practice. Eleven of the 335 surveys returned by current practitioners were excluded from the database due to varying degrees of incompleteness, which ranged

from blank surveys to the omission of entire sets of questions needed for statistical analyses. Thus, excluding unopened surveys that apparently never reached invited psychologists ($n = 126$), a total response rate of 19% (or 357 out of 1,874) was achieved with a survey return rate of 18% (345 out of 1,874). Of note, 94% of returned surveys (or 324 out of 345) contained usable data. Hence, the total return rate for usable surveys was 17% (or 324 out of 1,874).

Participants included in this study were 324 licensed psychologists in New England engaged in at least one hour of applied clinical practice (i.e., assessment or treatment) weekly throughout the past year (range = 1 to 100 weekly hours; $M = 24.84$, $SD = 12.41$). Numbers of participating psychologists by state were as follows: Massachusetts, $n = 135$ (42%); Connecticut, $n = 63$ (19%); Rhode Island, $n = 51$ (16%); Vermont, $n = 31$ (9%); New Hampshire, $n = 25$ (8%); and Maine, $n = 19$ (6%). Total years of post-graduate clinical service ranged from 3 to 57 ($M = 21.91$, $SD = 11.06$). The mean age of respondents was 55.36 years ($SD = 12.80$) with females ($n = 189$) making up the majority of the sample (58%). The vast majority (95%) of participants identified as Caucasian, 2% did not disclose their race, 1% identified as Asian, less than 1% as Black or African American, less than 1% as American Indian or Alaskan Native, and 1% as Hispanic or Latino. Nearly all (99%) of participants spoke English as a first language, and 94% of participants were born in the United States.

Regarding credentials, the majority of the sample (69%) reported having earned a Ph.D., 24% reported holding a Psy.D., 7% reported holding an Ed.D., and one participant reported having a Doctor of Ministry (D.Min.) degree. Most (78%)

doctoral degrees were in clinical psychology, followed by counseling psychology (15%), school psychology (4%), educational psychology (2%), and neuropsychology (1%). Ninety percent of the sample indicated that their graduate programs were APA-accredited, and nearly all of the remaining 10% reported Council for Accreditation of Counseling and Related Educational Programs (CACREP) accreditation. With reference to clinical settings in which respondents reported conducting most of their applied work, private or independent practice was most heavily represented (68%), and a cognitive-behavioral theoretical approach was most frequently endorsed (42%) within the sample (see Table 1 for a more detailed summary of demographic and professional characteristics of participants).

Measures

Clinical Attitudes and Knowledge Questionnaire (CAKQ). This newly developed questionnaire, the first of its kind to the best of the author's knowledge, was designed to quantify the nature of clinical knowledge and belief among licensed mental health practitioners. The initial version of the CAKQ consisted of 38 questions rated along a five-point Likert-type scale where 1 = "Strongly Disagree," 3 = "Neither Agree Nor Disagree," and 5 = "Strongly Agree." In addition, a response option of "Not Familiar" (NF) was provided in the event that respondents were unfamiliar with specific topics covered by items. Question content was relevant to applied clinical practice and covered five general themes, including knowledge about the research status of evidence-based treatment modalities (eight questions, e.g., "Psychological research has established cognitive-behavioral therapy [CBT] as an efficacious intervention for social anxiety disorder"), knowledge about the status of interventions

deemed pseudoscientific in the current literature (eight questions, e.g., “Past life regression is useful for identifying clients’ traumatic memories prior to their birth”), degree of familiarity with relevant areas of the broader clinical research literature (nine questions, e.g., “Psychotherapists, even when well intentioned, may unknowingly mislead clients into integrating completely fabricated events [i.e., false memories] into their personal histories”), knowledge about clinical judgment and decision making (eight questions, e.g., “Maximizing the accuracy of clinical judgments depends on integrating most or all of the available data”), and attitudes toward science in clinical psychology (five items, e.g., “Most of the soundest bases for knowledge in psychology rest on scientific studies and advances”).

Comprehensive edited volumes focusing on evidence-based (e.g., Chambless et al., 1998; Chambless & Ollendick, 2001; Lambert, 2013) and pseudoscientific clinical practices (e.g., Faust, 2012; Lilienfeld et al., 2003; Lilienfeld, Lynn, Ruscio, & Beyerstein, 2010; Lilienfeld et al., 2008) as well as a review of the general literature on these topics provided the basis for item content and design. Knowledge about the research status of clinical decision-making practices was considered equally important in this context given the pervasiveness of false beliefs about clinical versus actuarial judgment and potential repercussions for clinical practice (see Dawes, 1994; Dawes et al., 1989; Meehl, 1954). Specific treatment techniques were included based on persistence of use among practitioners (despite lack of evidence and/or contrary evidence for their effectiveness for specified purposes) and apparent public popularity according to the literature. For example, approximately 25% of American and Canadian psychotherapists regularly use questionable methods like hypnosis and

dream interpretation (as cited in Lilienfeld et al., 2008, p. 25), and substantially more public attention is garnered by highly questionable clinical techniques, such as past-life regression, TFT, and rebirthing vis-à-vis more evidence-based practices (Olatunji, Parker, & Lohr, 2006).

Rational-Experiential Inventory (REI): Selected Items. The REI is a widely used questionnaire validated for differentiating rational and intuitive thinking styles (Epstein, Pacini, Denes-Raj, & Heier, 1996; Pacini & Epstein, 1999). All items are rated along a five-point Likert-type scale where 1 = “Definitely False,” 3 = “Undecided or Equally True and False,” and 5 = “Definitely True.” The REI contains two negligibly correlated scales: Rationality (formerly the Need for Cognition scale), which measures rational thinking preferences (e.g., “I prefer complex problems to simple problems”), and Experientiality (formerly the Faith in Intuition scale), which assesses more intuitive thinking styles (e.g., “I believe in trusting my hunches”). To reduce questionnaire length, the present study included five items from the Experiential Ability (EA; $\alpha = .80$) subscale of the Experientiality scale, which refers to perceiving oneself as having sophisticated intuitive skills (e.g., “I hardly ever go wrong when I listen to my deepest ‘gut feelings’ to find an answer”); and five items from the Experiential Engagement (EE; $\alpha = .79$) subscale of the Experientiality scale, which refers to a preference for and enjoyment of intuitive decision making (e.g., “I like to rely on my intuitive impressions”) (EA-EE subscale inter-correlation = .62). These items were selected on the basis of the magnitude of item-total correlations (range of $r_s = .47$ to $.73$; Björklund & Bäckström, 2008), magnitude of respective factor loadings (range = $.50$ to $.66$; Pacini & Epstein, 1999), and their perceived

relevance to intuitive clinical decision making. Questionnaire instructions were slightly modified to associate each item with clinical practice activities.

Critical Thinking Questionnaire (CTQ). The CTQ was designed to evaluate practicing psychotherapists' critical thinking abilities (Gaudiano, et al., 2011; 2012; Sharp et al., 2008). It consists of 28 items with multiple-choice format response scales and has a total score range of 0 to 28; total scores are calculated by summing the number of correct responses. Kuder-Richardson Formula 20 (*KR20*) estimates of internal consistency appear adequate ($KR20 = .70$; Sharp et al., 2008), although additional psychometric evaluation is needed. Of the five CTQ subscales, three were selected for the current study (viz., Inference, Deduction, and Interpretation, comprising 14 items total) in light of (a) Sharp et al.'s (2008) observation that Deduction (4 items) and Interpretation (7 items) best differentiated psychotherapists' critical thinking practices, (b) the high clinical relevance of the Inference questions (3 items), and (c) concerns about the overall length of the current questionnaire packet. Inference questions challenge respondents to discern degrees of accuracy of inferences drawn from available data, Deduction problems test the ability to ascertain whether certain conclusions necessarily follow from given premises, and Interpretation questions require drawing accurate conclusions and generalizations upon weighing available evidence (Sharp et al., 2008). Of note, most CTQ questions originally appeared in two well-validated critical thinking measures—the Watson-Glaser Critical Thinking Assessment (Watson & Glaser, 1994) and the Cornell Critical Thinking Test (Ennis, Millman, & Tomko, 1985). Two Inference subscale items were adapted from Stanovich's (2003) textbook on critical thinking.

Procedure

Prior to participant recruitment, the initial CAKQ item pool was distributed to four nationally recognized psychologists with scholarly expertise in the delineation of scientific from pseudoscientific subject matter in clinical psychology. All four experts provided feedback in the form of suggested item revisions, new items for consideration, and suggested item omissions. No explicit conflicting feedback was observed among experts, and virtually all of their suggestions were incorporated into the revised questionnaire with the close guidance of the author's mentor, who has a strong background in the decision-making literature. In accordance with expert feedback, three new items were added to the final version of the CAKQ (i.e., one item to the evidence-based treatment category, one item to the pseudoscientific intervention category, and one item to the clinical judgment category), and four items in the general clinical knowledge category pertinent to scientifically under-supported assessment practices (e.g., use of graphology and Draw-A-Person test) were eliminated and replaced with four questions reflecting more general content (e.g., "Most, if not all, major psychopathology ultimately has its roots in low self-esteem"). The latter change was made in light of the observation that potentially pseudoscientific assessment practices comprise a separate domain of inquiry that is equally as complex as the intervention topic, and studying clinician knowledge of such practices was deemed best reserved for a future study. Thus, the final version of the CAKQ contained 41 items.

In light of the reliance of previous research on online listservs (i.e., e-mail advertisements) and membership rosters associated with professional psychological

organizations (e.g., Gaudiano et al., 2011; 2012; Sharp et al., 2008), this study attempted to reduce sampling error by recruiting participants from the northeastern U.S. irrespective of professional membership status. New England was specifically chosen for recruitment due to the known effects of physical proximity and personalization on survey return rates (i.e., the location of University of Rhode Island [URI] in a New England state and the personal relevance of practicing psychology in the Northeast; see Green & Kvidahl, 1989; Heerwegh, Vanhove, Matthijs, & Loosveldt, 2005; Sonne-Holm, Sørensen, Jensen, & Schnohr, 1989). Additionally, only a few studies to date (e.g., Hipol & Deacon, 2012) have focused on clinicians within specific geographic regions of the U.S.

A sample of potential research participants was identified using computer databases obtained from all six New England state boards of professional licensure. The number of randomly selected participants from each state database was proportional to the documented number of licensed psychologists actively practicing in each state, specifically, 160 from Rhode Island, 1,040 from Massachusetts, 360 from Connecticut, 160 from Vermont, 160 from New Hampshire, and 120 from Maine ($N = 2,000$). Most databases were freely downloadable from state licensure board websites, although two states required formal written requests (viz., Rhode Island and Massachusetts), and one state required a fee (viz., Massachusetts). The expected response rate was approximately 20% (actual response rate = 19%). This estimation was based on postal survey response rates of licensed mental health professionals reported in the literature, with doctoral-level practitioners being among the more responsive participants (see Michalski & Kohout, 2011; Sharp et al., 2008).

An explanatory cover letter, URI Institutional Review Board (IRB)-approved informed consent form, paper questionnaire packet (see Appendix), and pre-stamped return envelope were sent to invited participants' mailing addresses via U.S. Post Office First-Class Mail. All returned packets were screened for degree of completion and adherence to inclusion criteria (e.g., checking weekly clinical hours and respondents' hand-written notes for indications of retirement). Questionnaire data were entered into an SPSS Statistics 22.0 database for statistical analyses (see Data Analysis section below for details).

Data Analysis

Data from 324 participants were used for all statistical analyses. All but five CAKQ items were included in the analyses ($n = 36$ items). The five excluded items comprised the attitudes toward science category, which elicited personal attitudes (versus research-corroborated knowledge) and were not included in the hypotheses of the current study. Prior to conducting the primary statistical analyses, 17 CAKQ items were recoded such that higher scores reflected greater knowledge of relevant clinical research, 4 REI items were recoded such that higher scores indicated a greater preference for intuitive decision making, and missing data frequencies were obtained for all study questionnaires (see Appendix for specific examples of recoded items). In addition, frequencies were obtained for NF responses on the CAKQ. NF responses were treated as missing data given that these responses were distinct from the CAKQ Likert scale options and thus were not accounted for by the metric of the provided response scale, and idiographic-level (i.e., participant-specific) NF rationales were not obtained in this study due to time and resource constraints.

The following procedures were subsequently used to evaluate whether NF responses in the CAKQ were best characterized as missing at random (MAR) or missing completely at random (MCAR)¹¹ (for technical details, see Enders, 2010; Rubin, 1976). First, a multiple regression analysis (MRA) was performed to ascertain whether demographic and professional background variables significantly predicted total NF responses across participants. Second, Little's (1988) MCAR test was conducted as an omnibus test of randomness to judge whether NF responses plus the small amount of non-NF missing data could be collectively classified as MCAR (i.e., comparing expected data patterns from a random missing data process to observed missing data patterns). Deciphering the degree of randomness present in missing data has important implications for judging whether data are sufficiently random to accommodate specific remedial techniques, such as maximum likelihood estimation in the event of non-randomness/MAR, or a family of missing data imputation approaches if MCAR holds (e.g., mean substitution, regression imputation, hot or cold deck imputation, and multiple imputation) (Little & Rubin, 1987).

Next, a preliminary investigation of the dimensionality of the CAKQ was conducted using principal components analysis (PCA). Of note, the final sample size was not sufficiently large to allow a random division into equivalent subsamples ($ns = 162$) for cross-validation purposes (e.g., using a combinatory exploratory and confirmatory factor analytic approach with structural equation modeling would be inadequately powered and requires larger sample sizes; cf. Brown, 2006). The PCA was conducted in SPSS 22.0 using oblique rotation. Scree test results, solution interpretability, and strength of parameter estimates (i.e., component loadings $> .30$)

guided component selection and acceptability. Salient cross-loadings (i.e., $> .30$) were also identified. Following the PCA, a standard MRA was used to test hypothesized relationships between the CAKQ total score (outcome variable) and the cognitive/information processing variables—the REI total scores and CTQ domains (predictor variables). Finally, a simple linear regression analysis was used to determine whether reported years of clinical experience significantly predicted CAKQ total scores.

CHAPTER 4

RESULTS

CAKQ Response Frequencies and Characteristics

Prior to recoding of CAKQ items, the Likert-scale response frequencies for each item were computed and are presented in Table 2. In addition, an alternative scoring procedure was carried out wherein responses to CAKQ items ($n = 36$) were coded either “correct” or “incorrect” in accordance with contemporary clinical research findings (see Chambless & Ollendick, 2001; Faust, 2012; Lambert, 2013; Lilienfeld et al., 2003 for reviews relevant to specific item content). Specifically, if participants (a) responded with either “Strongly Disagree” or “Somewhat Disagree” to scientifically unsubstantiated statements (e.g., “Past life regression is useful for identifying clients’ traumatic memories prior to their birth”); or (b) responded with either “Strongly Agree” or “Somewhat Agree” to scientifically supported statements (e.g., “Exposure plus response prevention [ERP] is an effective psychological treatment for obsessive-compulsive disorder [OCD]”), their responses were considered “correct” and coded as 1s. Responses in the incorrect direction (e.g., either strongly or somewhat agreeing or disagreeing with scientifically unsubstantiated or substantiated statements, respectively), responses of “Neither Agree Nor Disagree,” NF responses, and missing data were all counted as “incorrect” and coded as 0s. Total CAKQ scores were subsequently computed by summing the number of correct responses across the 36 items for each participant. The CAKQ total score binomial distribution appeared

mesokurtic ($G_2 = -.10$) with no substantial skewness ($G_1 = 0.20$) noted.

Using the alternative scoring system, the mean CAKQ total score was 17.84 ($SD = 4.87$) out of a total possible score of 36 (i.e., an average score of 50% correct). The most frequently correctly answered item was: “Psychological research has established cognitive-behavioral therapy [CBT] as an efficacious intervention for social anxiety disorder” (92% correct), and the most frequently incorrectly answered item was: “Maximizing the accuracy of clinical judgments depends on integrating most or all of the available data (e.g., clinical observations, test results, interview data, etc.)” (3% correct). Overall, participants evidenced disproportionately poorer performance on clinical judgment (i.e., range of correct responses = 3% to 58%) and pseudoscience items (range correct = 15% to 73%) compared to evidence-based treatment (range correct = 36% to 92%) and general clinical knowledge items (range correct = 12% to 83%). However, it is important to recognize that 4 items with some of the highest NF response rates (i.e., four pseudoscience items and one general clinical knowledge item; NF response range = 109 to 186) were represented among the 10 most frequently incorrectly answered items (range correct = 3% to 27%). Thus, counting NF responses as incorrect may have rendered this particular scoring system overly punitive, which is why alternative procedures for handling these responses were subsequently considered¹².

Missing Data and “Not Familiar” Response Analysis

Missing data frequencies were reviewed for the abbreviated CTQ ($n = 14$ items), the REI ($n = 10$ items), and for the CAKQ items included in the statistical analyses ($n = 36$). Across all participants, no CTQ responses were missing, 3 REI

responses (< 0.1%) were missing, and 21 CAKQ responses (0.2%) were missing. Thus, missing responses alone were not deemed a major concern.

NF responses were next analyzed for the CAKQ. NF responses comprised 15% of all CAKQ data collected, with 82% ($n = 267$) of all participants providing at least one NF response and 97% of all remaining CAKQ items ($n = 35$) having at least one NF response. Given this observation, a Pareto chart of the 36 CAKQ items (see Figure 1 for the bar chart and Table 2 for NF response frequencies for each item) was requested to gather additional descriptive information (i.e., to ascertain whether certain items elicited more NF responses relative to other items). Visual inspection of the Pareto chart revealed that items containing pseudoscientific content appeared to elicit disproportionately more NF responses relative to other items (i.e., NF frequency range = 41 to 187, or 45% of all NFs), whereas clinical judgment items garnered disproportionately fewer NF responses (i.e., NF range = 0 to 38, or 0.1% of all NFs). However, NF responses to EBT (range = 2 to 144, or 27% of all NFs) and general clinical knowledge (range = 4 to 133, or 19% of all NFs) items appeared more evenly distributed.

To explore the possibility of similar NF patterns emerging within specific sub-groups of participants, a standard MRA was conducted with total NF responses (i.e., summed across all items for each participant) as the criterion and theoretical orientation, degree field, highest degree earned, and total years of clinical experience as predictor variables¹³. The overall regression model did not significantly predict NF responses, and the effect size was negligible ($R^2 < .01$). This finding, in conjunction with the non-significant result obtained from Little's MCAR test, indicate that CAKQ

NF responses may be justifiably classified as sufficiently (albeit not necessarily *categorically*) MCAR (i.e., no distinguishable differences were detected across participants with differing NF response rates, at least insofar as can be deciphered from demographic and professional background variables measured in this study). Thus, imputation procedures were deemed appropriate for addressing NF responses, and a standard (i.e., ordinary least squares) regression imputation (i.e., conditional mean imputation; Buck, 1960)¹⁴ was utilized to address missing values and NF responses among the CAKQ data. Total scores were subsequently computed for regression-imputed CAKQ scores¹⁵.

With regard to an analysis of outliers among total NF scores, one case had a studentized residual of 3.26 (studentized deleted residual = 3.31), and seven cases evidenced Mahalanobis distances greater than 30 (D^2 values = 33 to 68). However, a review of these individual cases did not reveal any consistent features (e.g., relatively heterogeneous demographics, professional concentrations, years of experience, theoretical orientations, CTQ scores, and REI scores were observed). In addition, removing these cases and re-running the MRA did not yield meaningfully discrepant results.

Internal Consistency Estimates and Item-Total Correlations

Internal consistency estimates computed for the CAKQ (Likert-scale scores) and REI were adequate (Cronbach's α s = .76 and .88, respectively). Scale reliability for the CTQ, however, was poor ($KR20$ = .62) relative to previous estimates (e.g., Sharp et al., 2008), and correlations among subscales (i.e., Inference, Interpretation, and Deduction) were relatively small in magnitude (r s = .24 to .28). Corrected item-

total correlations (i.e., with individual item values subtracted from the total score) ranged from negligible to moderate ($r_s = .022$ to $.41$). For a complete listing of scale means, standard deviations, and intercorrelations, see Table 3.

Principal Components Analysis and Exploratory Factor Analysis

A PCA was performed on the CAKQ Likert-scale scores for the purposes of (a) seeking further justification for the use of a total score in subsequent analyses, and (b) preliminarily testing for the presence of the four hypothesized content categories (viz., pseudoscience, clinical judgment, evidence-based treatments, and general clinical knowledge). In light of the strength of observed CAKQ inter-scale and subscale-total score correlations (i.e., hypothesized scales moderately-to-strongly correlating with one another, $r_s = .20$ to $.60$, as well as with total scores, $r_s = .50$ to $.80$), an oblique (i.e., direct oblimin) rotation was selected. Scree test results indicated a possible two-component solution with observed eigenvalues as follows: 5.1, 2.4, 1.6, and 1.5 (variance explained = 14%, 7%, 5%, and 4%, respectively). However, the resulting pattern matrix revealed a largely uninterpretable solution as evidenced by inconsistent themes among within-component items (i.e., items 9, 13, 17, 20, 26, and 35 for component 1 and items 18, 20, 21, and 25 for component 2) as well as an abundance of salient (i.e., $> .30$) cross-loadings (range = $.31$ to $.72$). Based on these results, further examination of dimensionality was not pursued at this time, and the use of the CAKQ total score was deemed appropriate for subsequent analyses given the apparent absence of evidence for initially hypothesized CAKQ sub-domains within this sample.

An exploratory factor analysis (EFA) was next conducted on the 10-item REI

for the purposes of (a) investigating the appropriateness of using subscale or total scores in subsequent analyses, and (b) corroborating the two-component (i.e., EE and EA) solution observed in the literature (e.g., Pacini & Epstein, 1999) using principal axis factoring. Oblique rotation (viz., direct oblimin) was used given the strong observed EE-EA inter-scale correlation ($r = .70, p < .001$). Factor selection and acceptability were guided by the scree test, solution interpretability, and strength of parameter estimates (i.e., primary factor loadings $> .30$). Scree test results clearly suggested a two-factor solution. Eigenvalues for the unreduced correlation matrix were as follows: 4.8, 1.0, .77, and .71 (variance explained = 48%, 10%, 8%, and 7%, respectively). Primary factor loadings for all 10 items were well above .30 (range = .46 to .80), no salient cross-loadings emerged, and the inter-factor correlation was strong ($r = .71$). The pattern matrix indicated that all EE items (i.e., REI items 6–10) loaded onto the first factor, and all but one EA item (i.e., REI item 3, which also loaded onto the first factor) loaded onto the second factor (see Table 4 for item means, standard deviations, factor loadings, and communalities). In view of the high EE-EA inter-factor correlation and their strong associations with the REI total score ($r_s = .93$ and .90, respectively), the REI total score was used in subsequent analyses.

Multiple Regression Analyses

Two separate standard MRAs were conducted to determine how well intuitive preferences, critical thinking skills, and number of years of clinical experience predicted clinical knowledge. CAKQ Likert-scale total scores (with regression imputation applied to NF and missing data responses prior to summation) served as the outcome variable in both analyses. In the first analysis, predictor variables were

REI total scores and CTQ subscale scores (i.e., Inference, Interpretation, and Deduction), which significantly predicted CAKQ scores, $R^2 = .23$ (adjusted $R^2 = .22$), $F(4, 316) = 23.51, p < .001$. Both REI total and the Inference subscale of the CTQ made significant contributions to the regression model, $t(316) = -7.48, p < .001$ ($\beta = -.38$), and $t(316) = 3.01, p = .003$ ($\beta = .16$), respectively. However, the Interpretation and Deduction subscales of the CTQ were not significant predictors (see Table 5 for a summary of regression results). Finally, the second MRA revealed that years of clinical experience did not significantly predict CAKQ scores, and the effect size was negligible (i.e., $R^2 < .01$).

CHAPTER 5

DISCUSSION

This study served as a preliminary investigation of the clinical research knowledge base¹⁶ of licensed psychologists in applied practice in New England using a newly developed questionnaire. Knowledge of specific clinical domains was also examined in relation to critical thinking skills and information processing styles. Relevant statistical analyses indicated that the initial hypotheses were partially supported. The hypothesis that a four-component solution (i.e., knowledge of pseudoscientific treatments, evidence-based treatments, general/miscellaneous clinical research, and clinical decision making) would best summarize the CAKQ data was not supported by PCA results. Specifically, the substantive meaning of the two components accounting for the most variance (i.e., 21%) was unclear. Various items associated with pseudoscientific, evidence-based, and general knowledge content areas loaded onto the first component, and mixed items pertaining to evidence-based practices and general clinical knowledge loaded onto the second component.

Of note, although minor thematic consistency (i.e., memory-related phenomena) was detected among items 9 and 26, this was deemed insufficient for extrapolating a coherent conceptual understanding from the first component. A possible reason for items failing to load onto their respective anticipated heuristic categories is inconsistency of clinician knowledge within the proposed domains. For example, on average, most clinicians may have been fairly knowledgeable about

specific kinds of EBTs, but endorsed patterns of EBT knowledge were variable both within individual clinician response sets as well as across different clinicians in the sample. This particular hypothesis thus may be inappropriate for future CAKQ studies given the observed lack of knowledge cohesion within clinical content areas. These results notwithstanding, PCA results preliminarily supported the use of a CAKQ total summed score given the apparent unidimensional nature of overall questionnaire responses, although more sophisticated statistical strategies (e.g., Rasch modeling within an item response theory framework; Yu, Popp, DiGangi, & Jannasch-Pennell, 2007) are better suited for addressing this issue.

The hypothesized relationship between clinical knowledge and critical thinking skills was partially supported. That is, higher CAKQ scores (and thus greater familiarity with the current state of clinical research findings) were significantly associated with higher Inference, but not Interpretation or Deduction, subscale scores on the CTQ. The positive relationship between critical inferential thinking skills and better knowledge of research-supported and under-supported interventions is no surprise; the ability to draw warranted inferences when reviewing scientific literature obviously contributes to building an accurate clinical knowledge base. The high clinical relevance of the three CTQ Inference questions likely also contributed to the prediction of CAKQ scores. The Deduction and Interpretation questions, however, may have involved critical thinking skills less relevant to clinical knowledge building (e.g., perhaps deduction and interpretation are less commonly utilized when reviewing research articles, attending workshops, and/or reading textbooks, in which authors and speakers typically draw generalizations and conclusions *for* professional consumers).

Consistent with expectations, a lower reliance on intuitive thinking styles (as measured by the total REI score) was associated with greater clinical knowledge. This in part may reflect an incompatibility between overreliance on “gut-level” intuitive judgments and deference to and/or affinity for seeking out relevant scientific research findings. For example, more intuitively inclined practitioners may be more naturally likely to trust their initial subjective judgments about newly encountered treatments and clinical claims, which may dampen motivation to obtain updated knowledge from contemporary research. Finally, the hypothesis that total number of years of clinical experience would not predict higher clinical knowledge scores was also upheld. This finding is consistent with previous research indicating that amount of hands-on clinical experience is all too often negligibly (if at all) related to professional competency (e.g., Goldberg, 1968; Lilienfeld et al., 2003; Vollmer et al., 2013; Weck et al., 2011).

With regard to the CAKQ NF response distribution (see Figure 1), participants generally indicated low levels of familiarity with (or perhaps were less confident in providing clear responses to) items addressing dubious clinical treatments (e.g., TFT and NLP) in contrast to high levels of familiarity with items tapping clinical judgment processes. The former observation may be partly explained by lower mainstream clinical practitioner exposure to pseudoscientific interventions than previously surmised (e.g., Olatunji et al., 2006). That is, many practitioners may be ignoring such approaches (or simply not reading about or otherwise researching them) for a variety of reasons, such as skepticism, lack of interest, lack of perceived applicability to individual professional practices, and/or feeling as if they already have a grasp of a sufficient repertoire of helpful therapeutic techniques. Another possibility is that

many practitioners simply may not have been exposed to the pseudoscience literature given its relatively narrow niche in clinical psychology when compared to other widely known lines of scholarly inquiry (e.g., EBTs).

Professed familiarity with clinical judgment content (as judged by disproportionately fewer NF responses to these items relative to other items) is not surprising in light of the previously cited decision-making literature, which reveals consistent themes of practitioner overconfidence in empirically under-supported or suboptimal clinical decisional strategies. Current results suggest that self-perceived familiarity with clinical judgment content in particular is by no means a trustworthy proxy for accuracy of knowledge about clinical judgment, especially its well-researched limitations. For example, despite considerable contrary evidence (e.g., see Grove et al., 2000; Sawyer, 1966), 94% of the sample agreed with the following claim: “Maximizing the accuracy of clinical judgments depends on integrating most or all of the available data (e.g., clinical observations, test results, interview data, etc.).”

A number of study limitations warrant attention. First, regarding demographic characteristics, the sample was predominantly comprised of older (i.e., 64% between 51 and 99 years of age) Caucasian (95%) psychologists in private or independent practice (68%), which limits the generalizability of results across different age, racial, and professional groups. Additional scale validation efforts should attempt to draw from more demographically and professionally diverse samples. Second, the professional background portion of the demographics questionnaire did not include a question about primary professional activities (e.g., which specific assessments and/or

interventions do practitioners spend most of their time conducting and with which particular clinical populations?). It was thus not possible to address questions about the uniform relevance of CAKQ content across participants with different professional foci and specializations. A third limitation was the use of abbreviated cognitive measures due to concerns about overall questionnaire length. Observed attenuations of scale reliabilities were most likely attributable to the substantially shortened versions of the REI and CTQ, and a more comprehensive assessment of potentially relevant critical thinking skills (as captured by the 28-item CTQ in particular) was not possible.

Because volunteers comprised the entire sample, it is important to consider the possible impact of self-selection bias on results, which would limit generalizability. Specifically, participants who chose to complete the survey (versus non-responders) may have felt more strongly or confidently about certain topics covered by questionnaire items (e.g., stronger beliefs about or preferences for or against evidence-based interventions, intuitive clinical judgment, the perceived importance of critical thinking in clinical contexts, etc.). Psychologists with stronger scientific leanings, for instance, may have been more willing to participate in a research study. Of note, an examination of central tendency and distribution characteristics did not reveal substantially skewed results for questionnaire scores, although the mean REI total score was slightly negatively skewed ($M = 32$, $SD = 6.3$), indicating a slightly disproportionate preference for intuitive clinical decision-making among respondents. Returning for a moment to the observation that predominantly older psychologists in private or independent practice with 10 or more years of post-graduate clinical

experience submitted completed surveys, it is possible that these individuals felt that survey content was more professionally salient or relevant for them, and/or perhaps they had more time and professional flexibility to complete the surveys compared to other psychologists (e.g., younger early-to-mid-career psychologists in other professional settings). Thus, the results of this dissertation may under-represent the belief profiles of younger (e.g., < 50 years of age) psychologists with 10 or fewer years of clinical service in non-private practice career settings. Future alternative sampling methodologies (e.g., stratified sampling or possibly cluster sampling from randomly selected hospitals, community health centers, and university clinical psychology departments) may assist with clarifying statistical relationships within a more demographically and professionally diverse group of clinical practitioners.

A noteworthy statistical concern within this study was the use of regression imputation to address NF responses (15% of all CAKQ data collected), which were treated as missing data. First, on a general note, regression imputation suffers from the shortcomings of overestimating correlations and underestimating variances (Little & Rubin, 1987), which may have biased study results. Multiple imputation (MI), however, is considered less biased than regression and mean imputation, in part due to the introduction of random noise into computation procedures (Allison, 2003). However, MI is not without its challenges. It is a broad and highly technical approach to missing data requiring specialized software, and disagreements exist about how many imputations are necessary under which circumstances, with suggestions ranging from 2 to 510 imputations for obtaining “good” statistical results (Graham, Olchowski, & Gilreath, 2007; Hershberger & Fisher, 2003).

Second, on a conceptual note, NF responses are inherently ambiguous in the sense that they are qualitatively different from agreeing with, disagreeing with, or taking a neutral position on a knowledge claim. The NF response option might be used, for example, by an intellectually cautious practitioner with rigorous clinical science training who may have knowledge about a particular intervention, but may not be comfortable responding in a perceived definitive manner. However, this same response may be liberally utilized by a less scientifically conscientious practitioner who fails to keep abreast of relevant clinical research literature. Thus, an optimal approach to analyzing NF responses is not readily apparent (especially without having respondents indicate their basis for selecting this response option), which is partly why they were treated as missing data for the current study.

Possible avenues for future scholarly inquiry may include questions such as the following: How might educators in the field of clinical psychology assist practitioners in developing and maintaining a professional attitude of skeptical open-mindedness toward known, typically used, and newly encountered approaches to assessment and treatment? With regard to ethical self-monitoring strategies, how might practitioners identify and attenuate personal biases and tendencies to yield to emotionally compelling and/or seemingly plausible clinical approaches in the face of conflicting evidence for their value? And more generally, what role might future psycho-educational strategies play in increasing adherence to the more effective appraisal systems science offers (e.g., deference to the current state of the evidence insofar as the peer-reviewed research literature reveals) given our well-known cognitive limitations when left to our own devices (e.g., the pervasive limitations of “gut-level”

intuition when faced with complex and/or high-stakes clinical decisions affecting vulnerable clients [Dawes, 1994])?

Future research might also examine the presence and extent of clinicians' cognitive discipline during specially designed decision making tasks (e.g., degree of willingness to defer to research evidence over emotional convictions) as well as endorsed degrees of openness and imperviousness to evidence contrary to personally preferred or default modes of assessment and treatment. Given the many known human shortcomings in appraising evidence (Hastie & Dawes, 2010), reinforcement of the scientific method in clinical decision making in tandem with the capacity for informed recognition and rejection of largely ineffective clinical methodologies becomes a critical prophylactic in a field that continues to impact the lives of many individuals.

The current study has underscored the importance of continuing to study individual clinical practitioners' cognitive styles and characteristics, which appear to be underemphasized in the extant literature. As discussed earlier, a sanitized conceptual divorce of pseudoscientific content from the actions of individuals (which are in turn associated with cognitive and emotional styles, convictions, and proclivities) may not be a productive emphasis. The author respectfully disagrees, for example, with the apparently common position, which is observed even among stellar scholars in the field (e.g., McNally, 2003; Tolin, 2013, May 28), that the antidote to "bad science" is simply conducting more "good science." This move debatably courts the aforementioned compartmentalization strategy by predominantly emphasizing content dissemination. Its limitations may become more apparent by drawing the

following analogy: If drug company *A* knowingly produces a pill repeatedly verified to be inert at best or harmful at worst (i.e., “bad pills”), all the while steadfastly maintaining that their pills work despite mutually incompatible evidence, is the appropriate sole response of the neighboring drug company *B* to continue mass producing pills repeatedly demonstrated to be effective (i.e., “good pills”)? Although no one would disagree that good pills should continue to be improved upon and appropriately disseminated, it seems difficult to ignore the pernicious influence of drug company *A*, whose disseminative reach and public impact may not necessarily be checked by the circumscribed responsible actions of company *B*. In addition, framing this complex problem as a mere horserace between “bad science” and “good science” (or “good” or “bad” psychotherapy, assessment, pharmacotherapy, etc.) ignores the cognitive, emotional, and motivational machinery associated with the ongoing aggressive propagation of pseudoscience.

Bringing this analogy back to the present example of the ethical design, research, justification, practice, and eventual broad implementation of effective psychotherapy, it is unlikely that well-meaning, “good scientists” who conduct methodologically sound research in earnest (e.g., in Ivory Tower academia) will necessarily single-handedly deter or mitigate the influence of pseudoscience in society at large. In fact, despite the presence of excellent quality research and the availability of efficacious and effective treatments, pseudoscience has remained remarkably pervasive and popular in the public spheres (Lilienfeld, Lynn, & Lohr, 2003; Lilienfeld, Ruscio, & Lynn, 2008). It is here that the imposition of sanctions, despite the infrequency with which it is mentioned, would seem to merit careful consideration.

That is, should licensing boards and/or professional organizations such as the APA formally reprimand (or, in cases involving certain categories of legal damages, potentially withdraw licensure from) clinical practitioners who persist in using known pseudoscientific interventions demonstrated to be ineffective or harmful (e.g., see Lohr & Fowler, 2002, summer; Lowman, 2012)?

Other possible steps for combating the prevalence and influence of pseudoscience among applied psychologists include the following: (a) revamping and possibly standardizing current clinical psychology graduate-level curricula to include formal training in clinical versus actuarial judgment, cognitive biases relevant to clinical decision making, philosophy of science considerations (e.g., distinguishing scientific from unscientific approaches), and applied psychometrics in assessment; (b) developing lists of “psychotherapies to avoid” (p. 8) based on evaluative criteria consistently applied to treatment studies, which would be similar to existing lists of ESTs (Chambless & Hollon, 1998; Chambless & Ollendick, 2001); (c) organizing pseudoscience watchdog groups to respond to popular self-help material and outlandish claims in the media with the help of professional psychological organizations; and (d) setting more rigorous standards for continuing education (CE) credit courses necessary for maintaining licensure (Lilienfeld, 1998, fall; Lilienfeld, 2010; Lohr & Fowler, 2002, summer; McFall, 1991). With regard to point “a” above, some psychologists (e.g., McFall, 1991) have argued that rote acquisition of facts has lamentably displaced learning core principles of scientific thinking in clinical psychology programs, and the APA has been loath to address this problem for far too long. Lohr and Fowler (2002, summer) went so far as to suggest that the APA (and

other accrediting bodies) should actively refuse to accredit clinical graduate programs that fail to adhere to a more rigorous academic curriculum emphasizing critical thinking skills and applied knowledge domains imperative to ethical practice. On a similar note, Lilienfeld (2010) decried the unchecked freedom of clinical psychology departments to implement whatever training program philosophy and emphasis they see fit with impunity, calling this decisional latitude “a grievous error” (p. 286).

However, it is also important to recognize limitations introduced by the distance between didactic and associated practical applications (e.g., clear behavioral modeling of core principles with subsequent practice opportunities combined with corrective feedback). For example, a survey of 653 clinical psychology graduate students from 169 U.S. doctoral programs indicated that although 60% of participants reported receiving satisfactory training in integrating science and practice (within a scientist-practitioner training model), approximately one-third of the sample acknowledged a disconnect between scientific knowledge learned in graduate school and their applied clinical behaviors later in their careers (VanderVeen, Reddy, Veilleux, January, & DiLillo, 2012). Thus, as emphasized by recent research (viz., Berenbaum & Shoham, 2011; Kaslow et al., 2004), standardizing and streamlining graduate curricula in line with the scientific literature may best be combined with an emphasis on supervised demonstrations of effective practical clinical applications of knowledge gleaned from research. This applied educational strategy may assist with improving the quality of early professionals’ clinical training, which in turn may help check the influence of pseudoscience. Additionally, user-friendly treatment evaluation tools, such as the “Therapy Rating Scale” proposed by Worrall (1990) for parents and

educators of children with learning disabilities, may be developed and tailored to assist clinicians with vetting novel treatments that pique their interest. Combined with the more rigorous educational training suggestions noted above, the dissemination of ready-to-use intervention evaluation tools (e.g., worksheets enumerating “full-stop” and “warning flag” criteria associated with specific treatment package features) among practitioners may help reinforce sorely needed applied critical thinking skills as well as diminish the influence of pseudoscience in applied practice.

Finally, on the front lines of actively practicing clinicians, making better use of available computer software to enhance the quality and ubiquity of helpful treatment outcome feedback is encouraged. For example, a fairly recent online psychotherapy outcome tracking program, the Systemic Therapy Inventory of Change (STIC), assesses and tracks both client progress and therapeutic alliance factors (Pinsof, Goldsmith, & Latta, 2012). Programs like STIC could be utilized more broadly in applied practice settings to furnish readily available clinical outcome data in a systematized and time efficient manner. The provision of on-demand, analyzed, and compactly summarized individual client data (e.g., using graphical software) to clinicians each session may greatly improve treatment planning, client-therapist collaboration, and clinician accountability beyond what is gleaned from intuitive impressions of how clients are faring psychologically.

In closing, most people, including highly educated clinical psychologists, harbor misguided beliefs to various degrees. The objective of this project was not to ridicule, demonize, or ostracize psychologists who display erroneous clinical convictions and/or suboptimal critical thinking skills. Rather, the objective was to

assist with beginning to uncover possible variables that lead to misguided beliefs and attitudes among those providing professional care to psychologically vulnerable (and often disempowered) people who place their trust in the contemporary mental health care system. It is clear that these beliefs do not exist in a vacuum or in a strictly academic context; rather, they are rife in a profession where their influences may adversely impact clients' lives. Uncovering the variables undergirding pseudoscientific beliefs and proclivities may aid in the creation of future psycho-educational strategies for clinical psychologists, which may mitigate negative influences in a professional climate demanding accountability, ethical treatment, and professional responsibility.

As we are reminded by David Hume (1748/2007), and more contemporarily, Carl Sagan (1995), extraordinary claims require extraordinary evidence. Although clinical psychology requires a delicate balance of skepticism and open-mindedness like any other scientific endeavor (Lilienfeld, Lynn, & Lohr, 2003), we must not, as space journalist James E. Oberg once quipped, remain so open-minded that our "brains fall out" in the process (Sagan & Druyan, 1997). However, as eloquently articulated by the late Richard Feynman (1998), "...in life, in gaiety, in emotion, in human pleasures and pursuits, and in literature and so on, there is no need to be scientific" (p. 2). That is, one need not rigidly adhere to an all-pervasive, lockstep scientism across all facets of life, which would be nonsensical in many contexts (e.g., deciding which music to enjoy or which books to read for pleasure). Science can guide our efforts in constructing a laser and prolonging longevity, but as Hume's guillotine reminds us, it cannot necessarily tell us where we *ought* to aim the laser or

under what circumstances we *ought* to postpone death (Leahey & Leahey, 1983). It likewise would be inappropriate and overly stilted to allow scientific concerns to dictate every therapeutic micro-movement and interaction occurring within the walls of clinicians' offices. As metaphorically illustrated by the Yerkes-Dodson law and other U-shaped functions (e.g., Seery, 2011), an Aristotelian balance between excess (fanatical scientism) and deficiency (postmodernist rejection of science) may suffice in working toward a broader implementation of ethically responsible and humane clinical care. In keeping with Alan Blinder's (1987) wise admonition, here's to the cultivation of hard heads and soft hearts in clinical psychology¹⁷.

FOOTNOTES

¹ The subtitle of this section is a tribute to the title of Alan F. Chalmers' (1999) classic text on the monumental efforts of eminent philosophers of science to formulate generalized definitions of science, which have often fallen short (e.g., the problems of naïve falsificationism).

² For example, in the early 1990s, a typical graduate school applicant was four times more likely to be accepted into a clinical Psy.D. program without tuition remission (compared to a clinical Ph.D. program), and students were enrolling in practitioner-oriented programs at a rate of nearly three times higher than enrollment rates in scientist-practitioner programs (Norcross, Hanych, & Terranova, 1996). As of 2010, the ratio of enrolled students to total number of applications submitted to APA-accredited Ph.D. programs in the U.S. was 5% compared to 18% for accredited Psy.D. programs (Kohout & Wicherski, 2010, October).

³ Of note, use of placebo comparison groups in psychotherapy research testing novel treatments is arguably neglectful of participants in certain contexts. For example, if existing treatments are already known to exceed the effects of placebo, then those interventions (TAUs) may best serve as the comparison group in place of the placebo. This reasoning is consistent with the Edinburgh Revision to the 1964 Declaration of Helsinki (see World Medical Association, 2013).

⁴ Identification of scientifically plausible mechanisms of change embedded in treatment packages (e.g., specific treatment components directly responsible for allaying distress) is another invaluable contribution of clinical science (Laurenceau et

al., 2007). Although an inability to isolate individual change mechanisms does not preclude using interventions otherwise shown to work, as has been the case with many psychopharmacological agents, this process may ultimately help protect against the undue crediting of non-specific treatment effects in many cases (Baker et al., 2008; Beyerstein, 1997).

⁵ Beyerstein (1997) appeared to consider “statistical significance” sufficient in this context, but it goes without saying that other important quantitative information (e.g., effect sizes, confidence intervals, and statistical power considerations) weighs heavily on the multifaceted problem of “meaningful” differences. Moreover, a litany of both logical (e.g., invalid application of *modus tollens* to probabilistic conditions via modal “rejection” or “failure to reject” a null hypothesis) and statistical (e.g., instances of insufficient power to detect genuine differences or over-sensitive designs whose results are interpreted as genuine effects without regard to effect sizes) problems have plagued null hypothesis significance testing (often abbreviated NHST) since its inception, which beclouds popularly discussed notions of theory confirmation and disconfirmation (e.g., the glaring disconnect between NHST results and pinpointing specific problems with theories). It is beyond the scope of this project to discuss these complex issues in detail; the interested reader is advised to consult Harlow, Mulaik, and Steiger (1997) as well as Morrison and Henkel (2009) for a more thorough overview of the problem.

⁶ Pseudoscience typically has been differentiated from other questionable forms of non-science in the extant literature. For example, “junk science” (see Edens et al., 2012; Huber, 1993) usually refers to dubious expert witness claims

masquerading as informed scientific opinion in courtrooms, which is usually at odds with legal rules of evidence (e.g., Frye or Daubert standards depending on the state; Faigman & Monahan, 2009). “Quackery,” on the other hand, has been defined by the U.S. Food and Drug Administration as “the commercialization of unproven, often worthless, and sometimes dangerous health products and procedures” (Young, 1988, p. 12).

⁷ An illustrative example of dismantling in the field of ethnobotany can be found in Wade Davis’ (1985) popular book (and later, Hollywood horror film) *The Serpent and the Rainbow*. In this book, Davis detailed his infiltration into Haitian voodoo secret societies to research an allegedly magical and nefarious “zombi powder” used by local sorcerers to exact revenge on enemies. When ingested in food or drink, the powder supposedly caused people to die and return from the grave as zombies, to which local Haitian village habitants attested. However, upon further scrutiny, Davis (1983) discovered that the voodoo sorcerer’s concoction contained a potent neurotoxin found in puffer fish called tetrodotoxin (TTX), which is known to cause a form of flaccid paralysis that gives the temporary appearance of death. Thus, although the superstitious zombification claim was debunked, Davis’ (1983) component analysis (or “dismantling”) style logic was noteworthy, for he demonstrated that the TTX alone (and not the human bone powder, tarantula innards, herbs, jimson weed, voodoo incantations whispered over the powder, etc.) was responsible for the observed temporary paralysis mistaken for death.

⁸ Some studies indicate that clinical experience sometimes may be negatively related to judgment accuracy. For example a study by Hermann and colleagues

(1999), found that older psychiatrists with more years of clinical experience adhered to outdated knowledge about electroconvulsive therapy and were unaware of updated research on this subject.

⁹ Mathematician Morris Kline's (1967) apt summary statement definitely applies here, viz., "human nature is a more complicated structure than a mass sliding down an inclined plane or a bob vibrating on a spring" (p. 499). Much of psychological research relies on intangible theoretical concepts inferred through quantification procedures yielding *latent constructs* or *latent variables* (or, in statistical terms, linear composites; for detailed statistical treatments, see Bentler, 1980; Bollen, 2002; and Loehlin, 2004). Latent variables remain widely accepted in contemporary psychological research and are ubiquitous across theories and explanatory frameworks (Borsboom, Mellenbergh, & van Heerden, 2003). Reification fallacies notwithstanding, these unobservable constructs (e.g., personality features/types, cognitive abilities, and emotional states) are typically viewed as giving rise to (or, in the author's diverging opinion, are numerical Platonic summary statements of) conglomerates of observable and measurable behaviors (e.g., social proximity), test scores (e.g., IQ scores), and patterns of physical arousal (e.g., physiological indicators such as blood pressure elevation). In the context of classical test theory, the unobserved value of a latent variable under a given set of circumstances is the *true score*, which we attempt to mathematically approximate as best we can despite the inevitability of varying degrees of distortion attributable to measurement error (cf. DeVellis, 2003, p. 15). Complicating matters further, invoking a scientific realist interpretation of factor analytic results, which some researchers

mistakenly do, we cannot declare with certainty that factors (or any of its constituent indicators, for that matter) share clean, isomorphic correspondences with their counterparts in reality, especially given the fragmentary, intangible, and uncontained nature of those counterparts (e.g., the diagnostic construct of depression as a finely fragmented array of various self-reported symptoms and observed behavioral signs).

¹⁰ Additional complicating factors include the difficulties often involved in reaching accurate conclusions as information at hand becomes increasingly complex and ambiguous, the salience and emotional seductiveness of privileging direct experience over more abstract research findings, and general human cognitive limitations. Given these considerations, erroneous conclusions, nonscientific leanings, and false beliefs arguably may be far more ubiquitous and influential than formally recognized by many psychologists (including Lilienfeld et al., 2003), perhaps even among those who otherwise perceive themselves as hardcore rationalists or scientists. Roughly speaking, clinicians are not exempt from being influenced by their own evolutionarily-shaped neurobiological wiring programmed to respond to immediate or near-immediate environmental circumstances with snap judgments. This reasoning is important to consider when one is tempted to attribute the false or otherwise unorthodox beliefs of clinicians to poor moral character and/or easily remedied “irrationality,” which may contribute to unjustifiably hasty and damaging characterological generalizations.

¹¹ Note that this terminology is superficially misleading. That is, MAR does not refer to haphazardly missing data. Rather, it refers to the presence of a systematic relationship between the likelihood of missing values (on Y) and other variables (X) in

a given dataset. In contrast, MCAR refers to the absence of a relationship between other variables and the probability of missing data on Y , which is closest to haphazardly missing data. Another missing data mechanism, termed missing not at random (MNAR), accounts for situations in which missing Y values are associated with Y values after controlling for the influence of other variables in a dataset (e.g., a disproportionate number of missing job performance ratings for workers who were fired for unacceptably low work performance after controlling for potential confounds, such as IQ) (Enders, 2010, pp. 5–8). Similar to MAR, there is no acknowledgement in the quantitative research literature of any formal evaluation procedure for clearly judging whether missing data are NMAR, and MNAR models (e.g., selection and pattern mixture models) rely on far more restrictive and generally untenable assumptions than MAR models (e.g., distribution assumptions that cannot be tested and controversial parameter estimation practices) (see Allison, 2002; Enders, 2010; Schafer & Graham, 2002).

¹² Of note, categorical data imputation was not conducted.

¹³ These particular independent variables were chosen based on perceived plausibility of accounting for variance in NF responses. For example, it is conceivable that practitioners with different theoretical orientation preferences, professional degrees/concentrations, and years of practice may have evidenced unique NF patterns.

¹⁴ Although not as quantitatively sophisticated as multiple imputation and maximum likelihood procedures, standard regression imputation is similar to these procedures in that it gleans information from non-missing data using complete-case analysis to construct regression equations used to impute missing values. It is

generally considered superior to mean substitution techniques (e.g., milder covariance attenuation), although correlations and multiple *R* values tend to be overestimated (Enders, 2010; Olinsky, Chen, & Harlow, 2003).

¹⁵ A comparison of results from imputed and non-imputed CAKQ data was omitted because the standard default approach to missing data (viz., the *complete case* or *listwise* approach) would have yielded misleading results (i.e., the sample size would have been drastically reduced from 324 to 59 participants, resulting in a substantial loss of statistical power).

¹⁶ Choosing between the terms “knowledge” and “belief” in this context is challenging, and one or the other term may not precisely characterize all CAKQ responses. This may depend in part on the specific wording of individual items, and it may also depend on idiographic characteristics of respondents, which become obfuscated with nomothetic analyses. Using the word “belief” in isolation is arguably an emotive and misleading term to some degree, invoking parallels to questions of “belief” in such things as the afterlife or the existence of life on other planets. Beliefs, of course, may be deemed warranted or unwarranted depending on the existence and nature of evidence bearing upon such questions. The purpose of developing the CAKQ was to detect practitioners’ familiarity with or knowledge of the current clinical literature, and many of the questions were explicitly worded to that end, for example, “Most scientific studies indicate that...” and “There is little to no scientific evidence that...”

On an idiographic level, there is no strict functional behavioral difference between clinician A, who does aromatherapy for OCD and has only a vague (or

possibly mistaken) idea about exposure and response prevention (ExRP) and its effectiveness, and clinician B, who knows precisely what ExRP is (and knows that the research shows it to be efficacious and effective for OCD, especially in conjunction with *Luvox*), yet administers aromatherapy for OCD regardless. However, if clinician A strongly disagreed with a statement about the effectiveness of ExRP for OCD framed as “Scientific research shows...,” then this response would not consist of pure belief at work—it would be undergirded by an ignorance of the relevant literature. If clinician B were to strongly disagree with the same statement, however, this would indicate a more belief-infused denial of the data as opposed to ignorance of the research.

¹⁷ A “hard-head, soft-heart” (cf. Blinder, 1987) approach to psychotherapy might draw from “science-informed humanism” (Allen, 2013, spring), which emphasizes both specific (e.g., evidence-based treatment components) and non-specific (e.g., empathy and therapeutic alliance) factors shown to be effective in treating psychopathology (Allen, 2013, spring; Bracken et al., 2012). This idea is compatible with Peterson’s (2000) concept of the *scientific practitioner*, although his admonition should be kept in mind: “Those who glorify the artistic aspects of clinical experience and resist scientific advances lead us astray” (p. 252).

Table 1

Demographic and Professional Characteristics of Participants ($N = 324$)

Characteristic	<i>n</i>	%
Age (years)		
30–40	53	16
41–50	62	19
51–60	66	21
61–70	118	36
71–80	21	6
81–90	2	<1
91–99	1	<1
Sex		
Female	189	58
Male	134	41
(<i>Missing data</i>)	1	<1
Race		
White or Caucasian	307	95
(<i>Missing data</i>)	6	2
Asian	5	1
Black or African American	3	<1
American Indian or Alaskan Native	3	<1
Highest degree earned		
Ph.D.	223	69
Psy.D.	78	24
Ed.D.	22	7
D.Min.	1	<1
Degree field		
Clinical psychology	253	78
Counseling psychology	47	15
School psychology	13	4
Educational psychology	5	2
Neuropsychology	3	1
Experimental psychology	1	<1

Religion and psychological studies	1	<1
Other psychology	1	<1
State of psychology licensure		
Massachusetts	135	42
Connecticut	63	19
Rhode Island	51	16
Vermont	31	9
New Hampshire	25	8
Maine	19	6
Theoretical orientation		
Cognitive-behavioral	135	42
Psychodynamic	68	21
Eclectic or integrative	35	11
Interpersonal	34	11
Systems or family	16	5
Humanistic	10	3
Strict behavioral or ABA	5	2
<i>(Missing data)</i>	5	2
Acceptance and commitment	4	1
Existential	3	1
Rational-emotive	2	<1
Feminist	2	<1
Clinical setting		
Private or independent practice	219	68
Hospital	25	8
University or four-year college	21	7
Other health service setting	15	5
Elementary or secondary school	14	4
Community health center	13	4
Government or VA medical center	9	3
Medical school	5	1
Correctional setting	2	<1
Other academic or research setting	1	<1
Hours of weekly clinical service		
1–10	50	15

11–20	85	26
21–30	104	32
31–40	71	22
41–50	12	4
51–100	2	<1
Years of clinical service		
1–10	56	17
11–20	101	31
21–30	89	27
31–40	67	21
41–50	9	3
51–100	2	<1

Table 2

Clinical Attitudes and Knowledge Questionnaire Item Responses

Item (Mean and Standard Deviation) and Response ($N = 324$)	n	%
1. Science certainly has its strong points, but its potential advantages over other methods for acquiring knowledge are often exaggerated or overperceived. ($M = 2.06$; $SD = 1.22$)		
NF = Not Familiar	1	<1
1 = Strongly Disagree	140	43
2 = Somewhat Disagree	97	30
3 = Neither Agree Nor Disagree	23	7
4 = Somewhat Agree	50	15
5 = Strongly Agree	13	4
2. Practicing clinicians do not usually develop high levels of insight into their clinical judgmental processes. ($M = 2.13$; $SD = 1.16$)		
NF = Not Familiar	6	2
1 = Strongly Disagree	104	32
2 = Somewhat Disagree	125	39
3 = Neither Agree Nor Disagree	29	9
4 = Somewhat Agree	51	16
5 = Strongly Agree	9	3
3. The majority of people who were sexually abused during childhood do not go on to develop severe personality disorders in adulthood. ($M = 3.89$; $SD = 5.46$)		
NF = Not Familiar	12	4
1 = Strongly Disagree	11	3
2 = Somewhat Disagree	44	14
3 = Neither Agree Nor Disagree	46	14
4 = Somewhat Agree	126	39
5 = Strongly Agree	84	26
Missing Data	1	<1
4. Empirically supported therapies rarely generalize to real-world settings. ($M = 2.09$; $SD = 1.05$)		
NF = Not Familiar	2	<1
1 = Strongly Disagree	102	32
2 = Somewhat Disagree	133	41
3 = Neither Agree Nor Disagree	44	14
4 = Somewhat Agree	37	11
5 = Strongly Agree	6	2
5. Psychological research has discredited the idea that human memory works like a video or tape recorder (e.g., that the brain is capable of near-perfect retention of the details of past events). ($M = 4.60$; $SD = 7.59$)		
NF = Not Familiar	21	7

1 = Strongly Disagree	8	3
2 = Somewhat Disagree	23	7
3 = Neither Agree Nor Disagree	16	5
4 = Somewhat Agree	81	25
5 = Strongly Agree	173	53
Missing Data	2	<1

6. Family therapy (e.g., the Maudsley model) is largely ineffective for treating anorexia nervosa. ($M = 1.22$; $SD = 1.27$)

NF = Not Familiar	144	44
1 = Strongly Disagree	40	12
2 = Somewhat Disagree	78	24
3 = Neither Agree Nor Disagree	53	16
4 = Somewhat Agree	6	2
5 = Strongly Agree	3	1

7. Acupuncture reduces symptoms of anxiety and depression by stimulating the flow of “qi” (energy) through bodily meridians. ($M = 2.09$; $SD = 1.75$)

NF = Not Familiar	110	34
1 = Strongly Disagree	25	8
2 = Somewhat Disagree	22	7
3 = Neither Agree Nor Disagree	78	24
4 = Somewhat Agree	70	22
5 = Strongly Agree	19	6

8. Nearly all clients who commit suicide have severe clinical depression. ($M = 2.73$; $SD = 1.32$)

NF = Not Familiar	9	3
1 = Strongly Disagree	46	14
2 = Somewhat Disagree	116	36
3 = Neither Agree Nor Disagree	35	11
4 = Somewhat Agree	89	28
5 = Strongly Agree	29	9

9. Hypnosis is **not** an effective tool for the accurate recovery of repressed memories (e.g., of past physical or sexual abuse). ($M = 3.18$; $SD = 1.64$)

NF = Not Familiar	41	13
1 = Strongly Disagree	9	3
2 = Somewhat Disagree	58	18
3 = Neither Agree Nor Disagree	41	13
4 = Somewhat Agree	94	29
5 = Strongly Agree	81	25

10. Psychotherapists may unknowingly mislead clients into integrating completely fabricated events (i.e., false memories) into their personal histories. ($M = 3.83$; $SD = 1.10$)

NF = Not Familiar	8	3
1 = Strongly Disagree	6	2

2 = Somewhat Disagree	23	7
3 = Neither Agree Nor Disagree	41	13
4 = Somewhat Agree	164	51
5 = Strongly Agree	82	25

11. Psychological research has established cognitive-behavioral therapy (CBT) as an efficacious intervention for social anxiety disorder. ($M = 4.40$; $SD = 1.01$)

NF = Not Familiar	6	2
1 = Strongly Disagree	8	3
2 = Somewhat Disagree	3	1
3 = Neither Agree Nor Disagree	8	3
4 = Somewhat Agree	108	33
5 = Strongly Agree	191	59

12. Directing most of one's attention to a client's uniqueness when making clinical predictions almost invariably decreases the overall level of judgmental accuracy. ($M = 3.36$; $SD = 9.36$)

NF = Not Familiar	38	12
1 = Strongly Disagree	41	13
2 = Somewhat Disagree	76	24
3 = Neither Agree Nor Disagree	81	25
4 = Somewhat Agree	69	21
5 = Strongly Agree	16	5
Missing Data	3	1

13. Past life regression is useful for identifying clients' traumatic memories prior to their birth. ($M = 1.14$; $SD = .94$)

NF = Not Familiar	57	18
1 = Strongly Disagree	211	65
2 = Somewhat Disagree	24	7
3 = Neither Agree Nor Disagree	20	6
4 = Somewhat Agree	9	3
5 = Strongly Agree	3	1

14. A large body of scientific research shows that clinical judgment is rarely superior to actuarial (or statistical) judgment methods in predicting outcomes (e.g., determining whether a client has a major depression, or whether someone will act violently). ($M = 3.42$; $SD = 7.70$)

NF = Not Familiar	35	11
1 = Strongly Disagree	29	9
2 = Somewhat Disagree	73	23
3 = Neither Agree Nor Disagree	56	17
4 = Somewhat Agree	79	24
5 = Strongly Agree	50	15
Missing Data	2	<1

15. Exposure-based interventions are **not** scientifically supported for treating post-traumatic stress disorder. ($M = 2.08$; $SD = 5.56$)

NF = Not Familiar	43	13
-------------------	----	----

1 = Strongly Disagree	113	35
2 = Somewhat Disagree	97	30
3 = Neither Agree Nor Disagree	24	7
4 = Somewhat Agree	33	10
5 = Strongly Agree	13	4
Missing Data	1	<1

16. The bilateral desensitization component (e.g., moving finger or light) of Eye Movement Desensitization and Reprocessing (EMDR) therapy effectively alleviates symptoms of posttraumatic stress disorder (PTSD). ($M = 3.17$; $SD = 5.59$)

NF = Not Familiar	52	16
1 = Strongly Disagree	27	8
2 = Somewhat Disagree	38	12
3 = Neither Agree Nor Disagree	39	12
4 = Somewhat Agree	127	39
5 = Strongly Agree	40	12
Missing Data	1	<1

17. Most scientific studies indicate that dreams accurately reflect unconscious autobiographical memories. ($M = 1.56$; $SD = 1.09$)

NF = Not Familiar	50	15
1 = Strongly Disagree	115	36
2 = Somewhat Disagree	108	33
3 = Neither Agree Nor Disagree	32	10
4 = Somewhat Agree	15	5
5 = Strongly Agree	4	1

18. Exposure plus response prevention (ERP) is an effective psychological treatment for obsessive-compulsive disorder (OCD). ($M = 4.06$; $SD = 5.56$)

NF = Not Familiar	44	14
1 = Strongly Disagree	3	1
2 = Somewhat Disagree	12	4
3 = Neither Agree Nor Disagree	19	6
4 = Somewhat Agree	92	28
5 = Strongly Agree	153	48
Missing Data	1	<1

19. There is little to no scientific evidence that neurolinguistic programming methods rapidly eliminate symptoms of specific phobias. ($M = 1.48$; $SD = 1.86$)

NF = Not Familiar	187	58
1 = Strongly Disagree	3	1
2 = Somewhat Disagree	23	7
3 = Neither Agree Nor Disagree	46	14
4 = Somewhat Agree	34	11
5 = Strongly Agree	30	10

20. The vast majority of clients diagnosed with DID have extensively corroborated histories of severe child abuse. ($M = 3.67$; $SD = 9.41$)

NF = Not Familiar	77	24
1 = Strongly Disagree	15	5
2 = Somewhat Disagree	25	8
3 = Neither Agree Nor Disagree	45	14
4 = Somewhat Agree	102	32
5 = Strongly Agree	57	18
Missing Data	3	1

21. Language and social skill deficits associated with autism can be treated effectively with early intensive behavioral intervention (EIBI; the Lovaas method). ($M = 2.54$; $SD = 1.91$)

NF = Not Familiar	101	31
1 = Strongly Disagree	11	3
2 = Somewhat Disagree	20	6
3 = Neither Agree Nor Disagree	32	10
4 = Somewhat Agree	124	38
5 = Strongly Agree	36	11

22. Most, if not all, major psychopathology ultimately has its roots in low self-esteem. ($M = 2.12$; $SD = 5.50$)

NF = Not Familiar	4	1
1 = Strongly Disagree	158	49
2 = Somewhat Disagree	87	27
3 = Neither Agree Nor Disagree	44	14
4 = Somewhat Agree	26	8
5 = Strongly Agree	4	1
Missing Data	1	<1

23. Tapping specific points on the body while thinking about a distressing problem — a component of Thought Field Therapy (TFT) — is not a well-supported intervention. ($M = 2.10$; $SD = 5.74$)

NF = Not Familiar	159	49
1 = Strongly Disagree	4	1
2 = Somewhat Disagree	31	10
3 = Neither Agree Nor Disagree	42	13
4 = Somewhat Agree	44	14
5 = Strongly Agree	43	13
Missing Data	1	<1

24. Clinical judgments and predictions of people's behaviors are often **not** much better than chance. ($M = 2.78$; $SD = 1.32$)

NF = Not Familiar	9	3
1 = Strongly Disagree	49	15
2 = Somewhat Disagree	99	31
3 = Neither Agree Nor Disagree	37	11
4 = Somewhat Agree	106	33
5 = Strongly Agree	24	7

25. Scientific research has refuted the notion that psychiatric hospital admissions are at their peak during full moons. ($M = 1.99$; $SD = 1.88$)

NF = Not Familiar	133	41
1 = Strongly Disagree	11	3
2 = Somewhat Disagree	32	10
3 = Neither Agree Nor Disagree	61	19
4 = Somewhat Agree	49	15
5 = Strongly Agree	38	12

26. Psychological research has consistently shown that recovered memory techniques are **not** trustworthy treatment methods for unearthing repressed memories. ($M = 3.43$; $SD = 1.68$)

NF = Not Familiar	49	15
1 = Strongly Disagree	3	1
2 = Somewhat Disagree	21	7
3 = Neither Agree Nor Disagree	31	10
4 = Somewhat Agree	128	40
5 = Strongly Agree	92	28

27. Vaccines contain toxins that may cause autism in young children. ($M = 1.65$; $SD = 5.49$)

NF = Not Familiar	17	5
1 = Strongly Disagree	228	70
2 = Somewhat Disagree	42	13
3 = Neither Agree Nor Disagree	21	7
4 = Somewhat Agree	14	4
5 = Strongly Agree	1	<1
Missing Data	1	<1

28. Attending to complex patterns and interrelationships among variables in clinical psychological test data is usually **unnecessary** for increasing the overall accuracy of clinical judgments. ($M = 2.10$; $SD = 5.51$)

NF = Not Familiar	20	6
1 = Strongly Disagree	128	40
2 = Somewhat Disagree	116	36
3 = Neither Agree Nor Disagree	25	8
4 = Somewhat Agree	25	8
5 = Strongly Agree	9	3
Missing Data	1	<1

29. Withholding critical incident stress debriefing (CISD) from people who have experienced a potentially traumatic event typically has harmful consequences. ($M = 2.60$; $SD = 5.62$)

NF = Not Familiar	69	21
1 = Strongly Disagree	42	13
2 = Somewhat Disagree	65	20
3 = Neither Agree Nor Disagree	45	14
4 = Somewhat Agree	73	23

5 = Strongly Agree	29	9
Missing Data	1	<1

30. We should respect different methods for trying to understand the world because other methods often equal or exceed science for acquiring knowledge. ($M = 3.53$; $SD = 7.64$)

NF = Not Familiar	6	2
1 = Strongly Disagree	42	13
2 = Somewhat Disagree	78	24
3 = Neither Agree Nor Disagree	71	22
4 = Somewhat Agree	89	28
5 = Strongly Agree	36	11
Missing Data	2	<1

31. Interpersonal psychotherapy (IPT) generally is **not** helpful for alleviating symptoms of major depressive disorder. ($M = 1.66$; $SD = 1.14$)

NF = Not Familiar	42	13
1 = Strongly Disagree	112	35
2 = Somewhat Disagree	119	37
3 = Neither Agree Nor Disagree	23	7
4 = Somewhat Agree	20	6
5 = Strongly Agree	8	3

32. Interoceptive exercises (e.g., deliberate shallow breathing exercises) are effective for treating panic symptoms. ($M = 3.24$; $SD = 5.65$)

NF = Not Familiar	69	21
1 = Strongly Disagree	23	7
2 = Somewhat Disagree	14	4
3 = Neither Agree Nor Disagree	33	10
4 = Somewhat Agree	118	36
5 = Strongly Agree	66	20
Missing Data	1	<1

33. In science, quantification offers powerful advantages over other ways of analyzing information or predicting outcomes. ($M = 3.66$; $SD = 5.49$)

NF = Not Familiar	26	8
1 = Strongly Disagree	10	3
2 = Somewhat Disagree	28	9
3 = Neither Agree Nor Disagree	74	23
4 = Somewhat Agree	127	39
5 = Strongly Agree	58	18
Missing Data	1	<1

34. Most of the soundest bases for knowledge in psychology rest on scientific studies and advances. ($M = 4.24$; $SD = 5.38$)

NF = Not Familiar	0	0
1 = Strongly Disagree	9	3
2 = Somewhat Disagree	30	9

3 = Neither Agree Nor Disagree	35	11
4 = Somewhat Agree	143	44
5 = Strongly Agree	106	33
Missing Data	1	<1

35. Research has shown that Dialectical Behavior Therapy (DBT) techniques reduce parasuicidal behaviors (i.e., non-suicidal self-injury) associated with borderline personality disorder (BPD). ($M = 4.03$; $SD = 1.39$)

NF = Not Familiar	28	9
1 = Strongly Disagree	1	<1
2 = Somewhat Disagree	2	<1
3 = Neither Agree Nor Disagree	16	5
4 = Somewhat Agree	131	40
5 = Strongly Agree	146	45

36. Maximizing the accuracy of clinical judgments depends on integrating most or all of the available data (e.g., clinical observations, test results, interview data, etc.). ($M = 4.58$; $SD = .79$)

NF = Not Familiar	2	<1
1 = Strongly Disagree	1	<1
2 = Somewhat Disagree	9	3
3 = Neither Agree Nor Disagree	9	3
4 = Somewhat Agree	77	24
5 = Strongly Agree	226	70

37. Much like any other way of trying to understand the world, scientific beliefs and conclusions are strongly, if not primarily, determined by such factors as culture and human biases. ($M = 3.19$; $SD = 1.22$)

NF = Not Familiar	1	<1
1 = Strongly Disagree	30	9
2 = Somewhat Disagree	80	25
3 = Neither Agree Nor Disagree	49	15
4 = Somewhat Agree	125	39
5 = Strongly Agree	39	12

38. Knowledge about the base rates (i.e., prevalence) of common mental disorders is very useful when making clinical diagnostic decisions. ($M = 3.43$; $SD = 1.19$)

NF = Not Familiar	0	0
1 = Strongly Disagree	23	7
2 = Somewhat Disagree	63	19
3 = Neither Agree Nor Disagree	50	15
4 = Somewhat Agree	129	40
5 = Strongly Agree	59	18

39. Research clearly indicates that the single greatest contributor to diagnostic and predictive accuracy in mental health is consistent adherence to one's clinical experience and judgments. ($M = 2.28$; $SD = 1.20$)

NF = Not Familiar	31	10
-------------------	----	----

1 = Strongly Disagree	40	12
2 = Somewhat Disagree	119	37
3 = Neither Agree Nor Disagree	81	25
4 = Somewhat Agree	46	14
5 = Strongly Agree	7	2

40. Nutritional approaches to treat ADHD symptoms lack consistent scientific support. ($M = 2.90$; $SD = 1.85$)

NF = Not Familiar	71	22
1 = Strongly Disagree	6	2
2 = Somewhat Disagree	51	16
3 = Neither Agree Nor Disagree	27	8
4 = Somewhat Agree	94	29
5 = Strongly Agree	75	23

41. If given access to identical information, practitioners with 10 or more years of clinical experience will tend to make considerably more accurate diagnostic judgments than practitioners with only a few years of clinical experience. ($M = 2.98$; $SD = 5.53$)

NF = Not Familiar	36	11
1 = Strongly Disagree	23	7
2 = Somewhat Disagree	79	24
3 = Neither Agree Nor Disagree	70	22
4 = Somewhat Agree	100	31
5 = Strongly Agree	15	5
Missing Data	1	<1

Table 3

Scale Means, Standard Deviations, and Intercorrelations

Name of Scale	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8
1. CAKQ Total	126.22	11.81	—	−.43**	−.40**	−.39**	.29**	.27**	.19**	.16**
2. REI Total	32.15	6.29		—	.90**	.93**	−.22**	−.20**	−.13*	−.16**
3. REI EA	16.95	3.08			—	.69**	−.23**	−.22**	−.14*	−.14*
4. REI EE	15.16	3.76				—	−.18**	−.16**	−.11	−.15**
5. CTQ Total	10.33	2.25					—	.70**	.81**	.58**
6. CTQ Inference	1.32	1.08						—	.25**	.24**
7. CTQ Interpretation	5.50	1.35							—	.28**
8. CTQ Deduction	3.51	.67								—

Note. CAKQ = Clinical Attitudes and Knowledge Questionnaire; REI = Rational-Experiential Inventory; EA = Experiential Ability; EE = Experiential Engagement; CTQ = Critical Thinking Questionnaire. CAKQ Total Scores were computed following regression imputation. * $p < .05$; ** $p < .01$ (two-tailed).

Table 4

Means, Standard Deviations, Rotated Factor Loadings, and Communalities for the Rational-Experiential Inventory Items

REI Item	<i>M</i>	<i>SD</i>	Factor loadings		<i>h</i> ²
			1	2	
3	3.99	.75	.46	.07	.26
6	3.03	.91	.58	.23	.58
7	2.69	.94	.62	.12	.50
8	3.21	1.01	.79	−.17	.47
9	2.95	1.04	.59	.16	.50
10	3.30	.93	.68	.10	.56
1	3.32	.92	.10	.67	.56
2	3.45	.85	.15	.64	.57
4	3.25	.88	.02	.48	.24
5	2.96	.88	−.06	.78	.55

Note. REI = Rational-Experiential Inventory. Boldface indicates highest factor loadings. REI item content can be found in the Appendix. Factor 1 = Experiential Ability; Factor 2 = Experiential Engagement; *h*² = communality.

Table 5

Multiple Regression Analysis Summary for Rational-Experiential Inventory Total Scores and Critical Thinking Questionnaire Subscale Scores Predicting Clinical Attitudes and Knowledge Questionnaire Total Scores

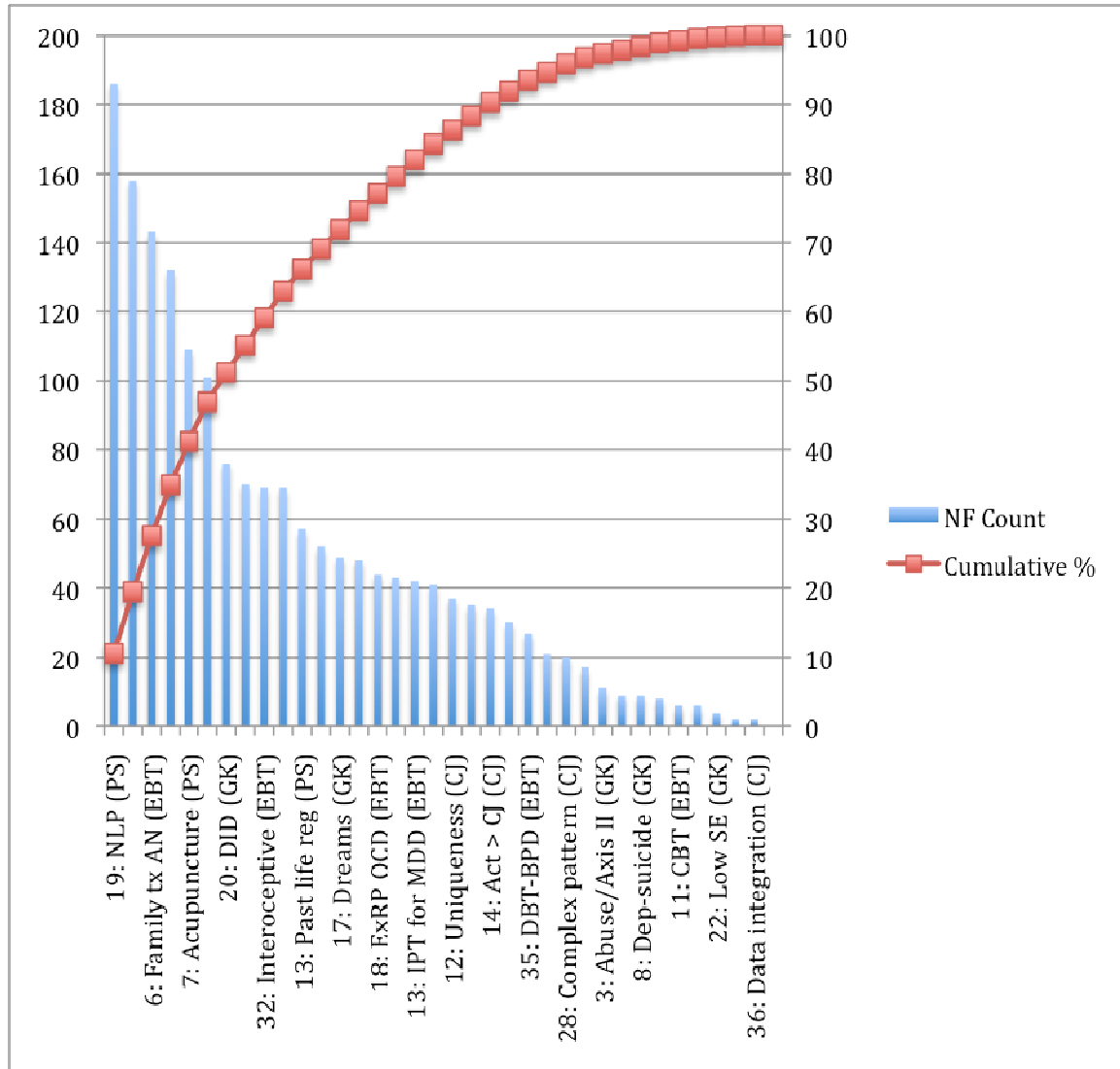
Predictor	<i>B</i>	<i>SE B</i>	β	95% CI	<i>t</i>	<i>p</i>
REI Total	−0.71	0.10	−.38	[−0.90, −0.53]	−7.48	< .001
CTQ Inference	1.73	0.57	.16	[0.60, 2.86]	3.01	.003
CTQ Interpretation	0.77	0.46	.09	[−0.13, 1.68]	1.69	.092
CTQ Deduction	0.77	0.92	.04	[−1.03, 2.58]	0.84	.400

Note. REI = Rational-Experiential Inventory; CTQ = Critical Thinking Questionnaire.

$R^2 = .23$ ($N = 321$, $p < .001$). CI = confidence interval for *B* (unstandardized regression coefficient).

Figure 1

Pareto Bar Graph of “Not Familiar” Responses to Clinical Attitudes and Knowledge Questionnaire Items



Note. On the x-axis, the number preceding the text descriptions is the CAKQ item number. NLP = Neurolinguistic Programming; PS = Pseudoscience item; AN = Anorexia Nervosa; EBT = Evidence-Based Treatment item; DID = Dissociative Identity Disorder; GK = General Clinical Knowledge item; ExRP = Exposure and response prevention; OCD = Obsessive-Compulsive Disorder; IPT = Interpersonal

Psychotherapy; MDD = Major Depressive Disorder; CJ = Clinical Judgment item; Act
= Actuarial prediction; DBT = Dialectical Behavior Therapy; BPD = Borderline
Personality Disorder; Dep = Depression; CBT = Cognitive-Behavioral Therapy; SE =
Self-Esteem.

APPENDIX

CLINICIAN SURVEY

PART A

Instructions: Please respond to the following items by checking the appropriate box or filling in the blank. Please do not check or write more than one answer for any item.

1. Age: _____
2. Sex: ☐ Male ☐ Female
3. Primary Language: ☐ English ☐ Other (please specify): _____
4. Race: ☐ American Indian/Alaskan Native ☐ Asian
☐ Native Hawaiian or Other Pacific Islander ☐ Black or African American
☐ White/Caucasian
5. Ethnicity: ☐ Hispanic or Latino ☐ Non-Hispanic or Latino
6. In what country were you born? ☐ United States ☐ Other (please specify): _____
7. Highest academic degree: ☐ Ph.D. ☐ Psy.D. ☐ Ed.D. ☐ Sc.D.
8. Field of highest academic degree: ☐ Clinical psychology ☐ Counseling psychology
☐ School psychology ☐ Clinical science ☐ Educational psychology
☐ Neuropsychology
9. Year highest degree obtained: _____
10. Did you graduate from an APA-accredited graduate psychology program?
☐ Yes ☐ No
11. Did you graduate from a CACREP-accredited graduate psychology program?
☐ Yes ☐ No
12. How many years have you provided clinical services (post-graduation)? _____

13. On average, about how many hours of clinical services have you provided weekly over the last year? _____

14. Setting in which you conduct the *highest percentage* of your current clinical practice:

- ☐ Private/independent practice ☐ Medical school ☐ University/four-year college
☐ Elementary or secondary school ☐ Community health center
☐ Government/VA medical center ☐ Correctional setting ☐ Hospital
☐ Other health service setting ☐ Other academic/research setting

15. What is your primary theoretical orientation? Please *check only one*.

- ☐ Cognitive-behavioral ☐ Strict behavioral/ABA ☐ Rational-emotive
☐ Psychodynamic ☐ Systems/family ☐ Interpersonal
☐ Acceptance & Commitment ☐ Humanistic ☐ Existential
☐ Other (please specify): _____

Instructions: The following items refer to various beliefs relevant to clinical psychology and the role of science. Using the 5-point response scale below, please indicate your level of agreement or disagreement with the items by writing the appropriate number on the line to the right of each item. Some of the diverse topics the items cover may be unfamiliar to you, in which case please write “NF” (for “Not Familiar”) on the line instead of a number.

1-----2-----3-----4-----5 or write **NF**
Strongly Somewhat Neither Somewhat Strongly
Disagree Disagree Agree Nor Agree Agree
Disagree

1. Science certainly has its strong points, but its potential advantages over other methods for acquiring knowledge are often exaggerated or overperceived. _____ (R) (SCI)
2. Practicing clinicians do **not** usually develop high levels of insight into their clinical judgmental processes. _____ (CJ)
3. The majority of people who were sexually abused during childhood do **not** go on to develop severe personality disorders in adulthood. _____ (GK)
4. Empirically supported therapies rarely generalize to real-world settings. _____ (R) (EBT)
5. Psychological research has discredited the idea that human memory works like a video or tape recorder (e.g., that the brain is capable of near-perfect retention of the details of past events). _____ (GK)

6. Family therapy (e.g., the Maudsley model) is largely ineffective for treating anorexia nervosa. _____ (R) (EBT)
7. Acupuncture reduces symptoms of anxiety and depression by stimulating the flow of “qi” (energy) through bodily meridians. _____ (R) (PS)
8. Nearly all clients who commit suicide have severe clinical depression. _____ (R) (GK)
9. Hypnosis is **not** an effective tool for the accurate recovery of repressed memories (e.g., of past physical or sexual abuse). _____ (PS)
10. Psychotherapists may unknowingly mislead clients into integrating completely fabricated events (i.e., false memories) into their personal histories. _____ (GK)
11. Psychological research has established cognitive-behavioral therapy (CBT) as an efficacious intervention for social anxiety disorder. _____ (EBT)
12. Directing most of one’s attention to a client’s uniqueness when making clinical predictions almost invariably decreases the overall level of judgmental accuracy. _____ (CJ)
13. Past life regression is useful for identifying clients’ traumatic memories prior to their birth. _____ (R) (PS)
14. A large body of scientific research shows that clinical judgment is rarely superior to actuarial (or statistical) judgment methods in predicting outcomes (e.g., determining whether a client has a major depression, or whether someone will act violently). _____ (CJ)
15. Exposure-based interventions are **not** scientifically supported for treating post-traumatic stress disorder. _____ (R) (EBT)
16. The bilateral desensitization component (e.g., moving finger or light) of Eye Movement Desensitization and Reprocessing (EMDR) therapy effectively alleviates symptoms of posttraumatic stress disorder (PTSD). _____ (R) (PS)
17. Most scientific studies indicate that dreams accurately reflect unconscious autobiographical memories. _____ (R) (GK)
18. Exposure plus response prevention (ERP) is an effective psychological treatment for obsessive-compulsive disorder (OCD). _____ (EBT)
19. There is little to no scientific evidence that neurolinguistic programming methods rapidly eliminate symptoms of specific phobias. _____ (PS)
20. The vast majority of clients diagnosed with DID have extensively corroborated histories of severe child abuse. _____ (R) (GK)
21. Language and social skill deficits associated with autism can be treated effectively

- with early intensive behavioral intervention (EIBI; the Lovaas method). _____ (EBT)
22. Most, if not all, major psychopathology ultimately has its roots in low self-esteem. _____ (R) (GK)
23. Tapping specific points on the body while thinking about a distressing problem — a component of Thought Field Therapy (TFT) — is not a well-supported intervention. _____ (PS)
24. Clinical judgments and predictions of people's behaviors are often **not** much better than chance. _____ (R) (CJ)
25. Scientific research has refuted the notion that psychiatric hospital admissions are at their peak during full moons. _____ (GK)
26. Psychological research has consistently shown that recovered memory techniques are **not** trustworthy treatment methods for unearthing repressed memories. _____ (PS)
27. Vaccines contain toxins that may cause autism in young children. _____ (R) (GK)
28. Attending to complex patterns and interrelationships among variables in clinical psychological test data is usually **unnecessary** for increasing the overall accuracy of clinical judgments. _____ (CJ)
29. Withholding critical incident stress debriefing (CISD) from people who have experienced a potentially traumatic event typically has harmful consequences. _____ (R) (PS)
30. We should respect different methods for trying to understand the world because other methods often equal or exceed science for acquiring knowledge. _____ (R) (SCI)
31. Interpersonal psychotherapy (IPT) generally is **not** helpful for alleviating symptoms of major depressive disorder. _____ (R) (EBT)
32. Interoceptive exercises (e.g., deliberate shallow breathing exercises) are effective for treating panic symptoms. _____ (EBT)
33. In science, quantification offers powerful advantages over other ways of analyzing information or predicting outcomes. _____ (SCI)
34. Most of the soundest bases for knowledge in psychology rest on scientific studies and advances. _____ (SCI)
35. Research has shown that Dialectical Behavior Therapy (DBT) techniques reduce parasuicidal behaviors (i.e., non-suicidal self-injury) associated with borderline personality disorder (BPD). _____ (EBT)

36. Maximizing the accuracy of clinical judgments depends on integrating most or all of the available data (e.g., clinical observations, test results, interview data, etc.). _____ (R) (CJ)
37. Much like any other way of trying to understand the world, scientific beliefs and conclusions are strongly, if not primarily, determined by such factors as culture and human biases. _____ (R) (SCI)
38. Knowledge about the base rates (i.e., prevalence) of common mental disorders is very useful when making clinical diagnostic decisions. _____ (CJ)
39. Research clearly indicates that the single greatest contributor to diagnostic and predictive accuracy in mental health is consistent adherence to one's clinical experience and judgments. _____ (R) (CJ)
40. Nutritional approaches to treat ADHD symptoms lack consistent scientific support. _____ (PS)
41. If given access to identical information, practitioners with 10 or more years of clinical experience will tend to make considerably more accurate diagnostic judgments than practitioners with only a few years of clinical experience. _____ (R) (CJ)

Note. Letters enclosed in parentheses following each item response blank were not printed on survey copies distributed to study participants and indicate the following: CJ = clinical judgment; EBT = evidence-based treatment; GK = general knowledge; PS = pseudoscience; R = reverse-scored item.

PART C

Instructions: Please rate the following statements about your feelings, beliefs, and behaviors **as they apply to your clinical practice activities** using the 5-point scale below. Again, please write the appropriate number on the line located to the right of each item.

1-----	2-----	3-----	4-----	5-----
Definitely	Mostly	Undecided or	Mostly	Definitely
False	False	Equally True	True	True
		and False		

1. If I were to rely on my gut feelings, I would often make mistakes. _____ (R) (EA)
2. I believe in trusting my hunches. _____ (EA)
3. I don't have a very good sense of intuition. _____ (R) (EA)
4. I trust my initial feelings about people. _____ (EA)
5. I hardly ever go wrong when I listen to my deepest "gut feelings" to find an answer. _____ (EA)
6. I like to rely on my intuitive impressions. _____ (EE)

7. I often go by my instincts when deciding on a course of action. _____ (EE)
8. I don't like situations in which I have to rely on intuition. _____ (R) (EE)
9. I don't think it is a good idea to rely on one's intuition for important decisions. _____ (R) (EE)
10. Intuition can be a very useful way to solve problems. _____ (EE)

Note. Letters enclosed in parentheses following each item response blank were not printed on survey copies distributed to study participants and indicate the following: EA = Experiential Ability; EE = Experiential Engagement; R = reverse-scored items.

PART D

Instructions: Please complete the following questions as best you can by darkening the boxed letter beside the correct answer. Remember that all responses are *completely confidential*. Please try to complete all questions in one sitting, do not spend too much time on any one question, and do not use help from others or other sources. Further instructions are provided for some of the exercises. Please make sure that all answers are clearly marked.

Exercise 1

Imagine that disorder X occurs in one in every 1,000 people. Imagine also there is a test to diagnose the disorder that always gives a positive result when a person has the disorder. Finally, imagine that the test has a false positive rate of 5 percent. This means that the test wrongly indicates that the disorder is present in 5 percent of the cases where the person does not have the disorder.

1. Imagine that we choose a person randomly, administer the test, and that it yields a positive result (indicates that the person has the disorder). What is the probability that the individual actually has the disorder, assuming that we know nothing else about the individual's psychological or medical history?

- ☐ A <10% ☐ B 10–30% ☐ C 30–50% ☐ D 50–70% ☐ E 70–90% ☐ F >90%
-

Exercise 2

The next exercises consist of brief paragraphs followed by several conclusions. For these questions, please assume that everything in the paragraph is true. The problem is to judge the whether or not each of the proposed conclusions logically follows beyond a reasonable doubt from the information given. Please mark either *follows* or *does not follow* after the conclusion.

Chris had poor posture, had very few friends, was ill at east around people, and in general was very unhappy. Then, a close friend recommended that Chris visit Dr. Carll, a reputed expert on helping people improve their personalities. Chris took this recommendation and, after three months of therapy with Dr. Carll, developed more friendships, was more at ease, and in general felt happier.

2. Without Dr. Carll's therapy, Chris would not have improved.

☐ A Follows ☐ B Does Not Follow

3. Without a friend's advice, Chris would not have heard of Dr. Carll.

☐ A Follows ☐ B Does Not Follow

When I go to bed at night, I usually fall asleep quite promptly. But about twice a month, I drink coffee during the evening, and whenever I do, I lie awake and toss for hours.

4. My problem is mostly psychological; I expect that the coffee will keep me awake, and therefore it does.

☐ A Follows ☐ B Does Not Follow

5. On nights when I want to fall asleep promptly, I'd better not drink coffee in the evening.

☐ A Follows ☐ B Does Not Follow

When the Journal Company, Inc. was created in 1960, it was the largest psychological journal company America had known up to that time. It produced twice as many psychological journals as all of its domestic competitors put together. Today, the Journal Company, Inc. produces about 20 percent of the psychological journals that are made in this country.

6. In 1960, the Journal Company, Inc. produced not less than 66 percent of the total domestic output of psychological journals.

☐ A Follows ☐ B Does Not Follow

7. Today, domestic competitors produce more than three times as many psychological journals as does the Journal Company, Inc.

☐ A Follows ☐ B Does Not Follow

8. The Journal Company, Inc. produces fewer psychological journals than it did in 1960.

☐ A Follows ☐ B Does Not Follow

Exercise 3

In this section, each exercise consists of several statements followed by several suggested conclusions. *For the purposes of this study, consider the statements in each exercise as true without exception.* After reading the conclusions beneath the statement, please mark whether you think it *follows* or *does not follow* from the statement given, regardless of whether you believe the statement to be true or not from your own experience or knowledge.

No person who thinks scientifically places any faith in the predictions of astrologers. Nevertheless, there are many people who rely on horoscopes provided by astrologers. Therefore—

9. People who lack confidence in horoscopes think scientifically.

☐ A Follows ☐ B Does Not Follow

10. Many people do not think scientifically.

☐ A Follows ☐ B Does Not Follow

Most persons who attempt to break their smoking habit find that it is something that they can accomplish only with difficulty, or cannot accomplish at all. Nevertheless, there is a growing number of individuals whose strong desire to stop smoking has enabled them to break the habit permanently. Therefore –

11. Only smokers who strongly desire to stop smoking will succeed in doing so.

☐ A Follows ☐ B Does Not Follow

12. A strong desire to stop smoking helps some people to permanently break the habit.

☐ A Follows ☐ B Does Not Follow

The table below summarizes data from an experiment:

	Improvement	No Improvement
<i>Treatment</i>	200	75
<i>No Treatment</i>	50	15

13. Based on these data, please rate the degree of effectiveness of the treatment on the following scale:

☐ A Not at all effective ☐ B Somewhat effective ☐ C Effective ☐ D Very effective

14. If the above table were an accurate reflection of the effectiveness of an innovative new treatment, how likely would you be to use it?

☐ A Would not use ☐ B Would possibly use ☐ C Would probably use

☐ D Would definitely use

BIBLIOGRAPHY

- Acierno, R., Hersen, M., Van Hasselt, V. B., Tremont, G., & Meuser, K. (1994). Review of the validation and dissemination of eye-movement desensitization and reprocessing: A scientific and ethical dilemma. *Clinical Psychology Review, 14*, 287–299.
- Addis, M. E., & Krasnow, A. D. (2000). A national survey of practicing psychologists' attitudes toward psychotherapy treatment manuals. *Journal of Consulting and Clinical Psychology, 68*(2), 331–339.
- Adúriz, M. E., Bluthgen, C., & Knopfler, C. (2009). Helping child flood victims using group EMDR intervention in Argentina: Treatment outcome and gender differences. *International Journal of Stress Management, 16*, 138–153.
- Aegisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., ...Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist, 34*(3), 341–382.
- Albright, D. L., & Thyer, B. (2010). Does EMDR reduce post-traumatic stress disorder symptomatology in combat veterans? *Behavioral Interventions, 25*, 1–19.
- Allen, J. G. (2013, spring). Psychotherapy is an ethical endeavor: Balancing science and humanism in clinical practice. *Bulletin of the Menninger Clinic, 77*(2), 103–131.
- Allison, P. D. (2002). *Missing data*. Newbury Park, CA: Sage.

- Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology, 112*(4), 545–557.
- Allumbaugh, D. L., & Hoyt, W. T. (1999). Effectiveness of grief therapy: A meta-analysis. *Journal of Counseling Psychology, 46*, 370–380.
- Altmaier, E. M. (2011). Best practices in counseling grief and loss: Finding benefit from trauma. *Journal of Mental Health Counseling, 33*(1), 33–45.
- American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*, 1060–1073.
- American Psychological Association, 2009 Presidential Task Force on the Future of Psychology as a STEM Discipline. (2010, June). *Psychology as a core science, technology, engineering, and mathematics (STEM) discipline*. Washington, DC: Author.
- Antony, M. M., & Barlow, D. H. (2002). Specific phobias. In: D.H. Barlow (Ed.), *Anxiety and its disorders: The nature and treatment of anxiety and panic* (2nd ed., pp. 380–417). New York: Guilford Press.
- Arkes, H. R. (1981). Impediments to accurate clinical judgment and possible ways to minimize their impact. *Journal of Personality and Social Psychology, 2*, 531–539.
- Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Science, 9*, 83–89.
- Baker, D. B., & Benjamin Jr., L. T. (2000). The affirmation of the scientist-practitioner: A look back at Boulder. *American Psychologist, 55*(2), 241–247.
- Baker, T. B., McFall, R. M., & Shoham, V. (2008). Current status and future

- prospects of clinical psychology: Toward a scientifically principled approach to mental and behavioral health care. *Psychological Science in the Public Interest*, 9(2), 67–103.
- Baskin, T. W., Tierney, S. C., Minami, T., & Wampold, B. E. (2003). Establishing specificity in psychotherapy: A meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology*, 71(6), 973–979.
- Bass, E., & Davis, L. (1990). *The courage to heal workbook: For women and men survivors of child sexual abuse*. New York: HarperCollins.
- Bausell, R. B. (2007). *Snake oil science: The truth about complementary and alternative medicine* (pp. 113–126). New York: Oxford University Press.
- Beecham, J., Hallam, A., Knapp, M., Baines, B., Fenyo, A., & Ashbury, M. (1997). Costing care in hospital and in the community. In J. Leff (Ed.), *Care in the community: Illusion or reality?* (pp. 93–108). Chichester, England: Wiley.
- Begley, S. (2009, October). Ignoring the evidence: Why do psychologists reject science? *Newsweek*, 154(15). Retrieved from <http://www.thedailybeast.com/newsweek/2009/10/01/ignoring-the-evidence.html>
- Belar, C., & Perry, N. W. (1992). National conference on scientist-practitioner education and training for the professional practice of psychology. *American Psychologist*, 47, 71–75.
- Bennett, J. C. (1991). Fundamentalists students: What place should they have in the public university? *College Student Journal*, 25, 515–517.

- Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, 107, 238–246.
- Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology*, 31, 419–456.
- Berenbaum, H., & Shoham, V. (2011). Broad and cutting-edge training in applied psychology: A clinical science perspective. *Training and Education in Professional Psychology*, 5(1), 22–29.
- Betan, E. J., & Binder, J. L. (2010). Clinical expertise in psychotherapy: How expert therapists use theory in generating case conceptualizations and interventions. *Journal of Contemporary Psychotherapy*, 40, 141–152.
- Beyerstein, B. L. (2001). Fringe psychotherapies: The public at risk. *The Scientific Review of Alternative Medicine*, 5(2), 70–79.
- Beyerstein, B. L. (1997). Why bogus therapies seem to work. *Skeptical Inquirer*, 21, 29–34.
- Bisson, J. L., Jenkins, P. L., Alexander, J., & Bannister, C. (1997). A randomized controlled trial of psychological debriefing for victims of acute harm. *British Journal of Psychiatry*, 171, 78–81.
- Björklund, F., & Bäckström, M. (2008). Individual differences in processing styles: Validity of the Rational-Experiential Inventory. *Scandinavian Journal of Psychology*, 49, 439–446.
- Blinder, A. S. (1987). *Hard heads, soft hearts: Tough-minded economics for a just society*. New York: Basic Books.
- Boghossian, P. (2007). *Fear of knowledge: Against relativism and constructivism*.

New York: Oxford University Press.

- Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, 53, 605–634.
- Bolognesi, F., Baldwin, D. S., & Ruini, C. (2014). Psychological interventions in the treatment of generalized anxiety disorder: A structured review. *Journal of Psychopathology*, 20(2), 111–126.
- Bolton, D. (2008). The epistemology of randomized, controlled trials and application in psychiatry. *Philosophy, Psychiatry, & Psychology*, 15(2), 159–165.
- Bonanno, G. A., & Lilienfeld, S. O. (2008). Let's be realistic: When grief counseling is effective and when it's not. *Professional Psychology: Research and Practice*, 39(3), 377–380.
- Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science*, 8(4), 445–454.
- Borkovec, T. D., & Castonguay, L. G. (1998). What is the scientific meaning of “empirically supported therapy”? *Journal of Consulting and Clinical Psychology*, 66, 136–142.
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, 110(2), 203–219.
- Boudewyns, P. A., Stwertka, S. A., Hyer, L. A., Albrecht, J. W., & Sperr, E. V. (1993). Eye movement desensitization for PTSD of combat: A treatment outcome pilot study. *Behavior Therapist*, 16, 29–33.

- Bower, B. (1995). Promise and dissent. *Science News*, 148, 270–271.
- Bracken, P., Thomas, P., Timimi, S., Asen, E., Behr, G., Beuster, C., ...Yeomans, D. (2012). Psychiatry beyond the current paradigm. *British Journal of Psychiatry*, 201(6), 430–434.
- Brown, H. (2013, March 25). Looking for evidence that therapy works. *The New York Times*. Retrieved from <http://www.nytimes.com>
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
- Buck, S. F. (1960). A method of estimation of missing values in multivariate data suitable for use with an electronic computer. *Journal of the Royal Statistical Society, Series B*, 22, 302–306.
- Bullis, J. R., Fortune, M. R., Farchione, T. J., & Barlow, D. H. (2014). A preliminary investigation of the long-term outcome of the unified protocol for transdiagnostic treatment of emotional disorders. *Comprehensive Psychiatry*. Advance online publication.
- Bunge, M. (1984, fall). What is pseudoscience? *Skeptical Inquirer*, 9, 36–46.
- Cahill, S. P., Carrigan, M. H., & Frueh, B. C. (1999). Does EMDR work? And if so, why?: A critical review of controlled outcome and dismantling research. *Journal of Anxiety Disorders*, 13, 5–33.
- Campbell, D. T. (1974). Evolutionary epistemology. In P. A. Schlipp (Ed.), *The philosophy of Karl R. Popper* (pp. 412–463). LaSalle, IL: Open Court.
- Campbell, D. T. (1956). Leadership and its effects upon the group. *Bureau of Business Research: Research Monograph*, 83, xi–92. Columbus: Ohio State

University.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Company.

Carey, S. S. (2011). *A beginner's guide to scientific method* (4th ed.). Boston: Wadsworth.

Carnap, R. (2003). *The logical structure of the world and pseudoproblems in philosophy* (R. A. George, Trans.). Peru, IL: Open Court.

Castonguay, L. G., Boswell, J. F., Constantino, M. J., Goldfried, M. R., & Hill, C. E. (2010). Training implications of harmful effects of psychological treatments. *American Psychologist*, 65(1), 34–49.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245–276.

Cautin, R. L. (2008). David Shakow and schizophrenia research at Worcester State Hospital: The roots of the scientist-practitioner model. *Journal of the History of the Behavioral Sciences*, 44(3), 219–237.

Cautin, R. L. (2011). Invoking history to teach about the scientist-practitioner gap. *History of Psychology*, 14(2), 197–203.

Chalmers, A. F. (1999). *What is this thing called science?* (3rd ed.). Queensland, Australia: University of Queensland Press.

Chambless, D., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Cristoph, P., ... Woody, S. R. (1998). Update on empirically validated therapies, II. *The Clinical Psychologist*, 51, 3–16.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies.

- Journal of Consulting and Clinical Psychology*, 66(1), 7–18.
- Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, 52, 685–716.
- Chambless, D. L., Sanderson, W. C., Shoham, V., Johnson, S. B., Pope, K. S., Crits-Cristoph, P., ...McCurry, S. (1996). An update on empirically validated therapies. *The Clinical Psychologist*, 49, 5–18.
- Chen, T. C., Christofferson, D. E., Hamlett-Berry, K., Kazerooni, R., Bounthavong, M., ...Myers, M. (2014). A tobacco cessation treatment model using telehealth: A pilot evaluation in Veterans. *Journal of Telemedicine and Telecare*, 20(3), 161–163.
- Committee on Crossing the Quality Chasm: Adaptation to Mental Health and Addictive Disorders. (2006). *Improving the quality of health care for mental and substance-use conditions*. Washington, DC: National Academies Press.
- Cooper, Z., & Shafran, R. (2008). Cognitive behaviour therapy for eating disorders. *Behavioural and Cognitive Psychotherapy*, 36(6), 713–722.
- Crits-Cristoph, P., Frank, E., Chambless, D. L., Brody, C., & Karp, J. F. (1995). Training in empirically validated treatments: What are clinical psychology students learning? *Professional Psychology: Research and Practice*, 26, 514–522.
- Davidson, P. R., & Parker, K. C. H. (2001). Eye movement desensitization and reprocessing (EMDR): A meta-analysis. *Journal of Consulting and Clinical Psychology*, 69, 305–316.

- Davis, W. (1983). The ethnobiology of the Haitian zombie. *Journal of Ethnopharmacology*, 23, 121–126.
- Davis, W. (1985). *The serpent and the rainbow*. New York: Simon & Schuster.
- Dawes, R. M. (1971). A case study of graduate admissions: Application of three principles of human decision making. *American Psychologist*, 26, 180–188.
- Dawes, R. M. (1989). Experience and validity of clinical judgment: The illusory correlation. *Behavioral Sciences & the Law*, 7(4), 457–467.
- Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York: Simon & Schuster.
- Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668–1674.
- Dazzi, C., & Pedrabissi, L. (2009). Graphology and personality: An empirical study on validity of handwriting analysis. *Psychological Reports*, 105(3), 1255–1268.
- de Dreu, C. K. W., & van Knippenberg, D. (2005). The possessive self as a barrier to conflict resolution: Effects of mere ownership, process accountability, and self-concept clarity on competitive cognitions and behavior. *Journal of Personality and Social Psychology*, 89(3), 345–357.
- de Jongh, A., Broeke, E. T., & Renssen, M. R. (1999). Treatment of specific phobias with eye movement desensitization and reprocessing (EMDR): Protocol, empirical status and conceptual issues. *Journal of Anxiety Disorders*, 13, 69–85.
- Department of Veterans Affairs & Department of Defense. (2004). *Clinical practice*

- guideline for the management of post-traumatic stress*. Washington, DC: Veterans Health Administration, Department of Veterans Affairs and Health Affairs, Department of Defense. Office of Quality and Performance publication 10Q-CPG/PTSD-04.
- Derksen, A. A. (1993). The seven sins of pseudo-science. *Journal for General Philosophy of Science*, 24(1), 17–42.
- Derry, C. J., Derry, S., McQuay, H. J., & Moore, R. A. (2006). Systematic review of systematic reviews of acupuncture published 1996–2005. *Clinical Medicine: Journal of the Royal College of Physicians of London*, 6, 381–386.
- DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). *Applied Social Research Methods Series* (Vol. 26). Thousand Oaks, CA: Sage.
- Deville, G. J. (2005). Power therapies and possible threats to the science of psychology and psychiatry. *Australian and New Zealand Journal of Psychiatry*, 39, 437–445.
- Deville, G. J., Gist, R., & Cotton, P. (2006). Ready! Fire! Aim! The status of psychological debriefing and therapeutic interventions: In the work place and after disasters. *Review of General Psychology*, 10(4), 318–345.
- Deville, G. J., & Lohr, J. M. (2008). Science and pseudoscience in victims' services. *International Review of Victimology*, 15(2), 105–122.
- Deville, G. J., & Spence, S. H. (1999). The relative efficacy and treatment distress of EMDR and a cognitive behavioral trauma treatment protocol in the amelioration of post traumatic stress disorder. *Journal of Anxiety Disorders*, 13, 131–157.

- Deville, G. J., Spence, S., & Rapee, R. (1998). Statistical and reliable change with eye movement desensitization and reprocessing: Treating trauma within a veteran population. *Behavior Therapy*, 26, 435–455.
- Donn, J. E., Routh, D. K., & Lunt, I. (2000). From Leipzig to Luxembourg (via Boulder and Vail): A history of clinical psychology training in Europe and the United States. *Professional Psychology: Research and Practice*, 31(4), 423–428.
- Dougherty, M. J. (2004). Educating believers: Research demonstrates that courses in skepticism can effectively decrease belief in the paranormal. *Skeptic*, 10(4), 31–35.
- Dimidjian, S., & Hollon, S. D. (2010). How would we know if psychotherapy were harmful? *American Psychologist*, 65(1), 21–33.
- Dunbar, K., & Fugelsang, J. (2005). Scientific thinking and reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 705–726). New York: Cambridge University Press.
- Duncan, D. F., Donnelly, J. W., Nicholson, T., & Hees, A. J. (1992). Cultural diversity, superstitions, and pseudoscientific beliefs among allied health students. *College Student Journal*, 26(4), 525–530.
- Edens, J. F., Smith, S. T., Magyar, M. S., Mullen, K., Pitta, A., & Petrila, J. (2012). “Hired guns,” “charlatans,” and their “voodoo psychobabble”: Case law references to various forms of perceived bias among mental health expert witnesses. *Psychological Services*, 9(3), 259–271.
- Einhorn, H. J. (1986). Accepting error to make less error. *Journal of Personality*

Assessment, 50, 387–395.

- Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.
- Ennis, R. H., Millman, J., & Tomko, T. N. (1985). *Cornell Critical Thinking Tests Level X & Level Z Manual* (3rd ed.). Pacific Grove, CA: Critical Thinking Books and Software.
- Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytic-rational thinking styles. *Journal of Personality and Social Psychology*, 71, 390–405.
- Ericsson, K. A. (2005). Recent advances in expertise research: A commentary on the contributions to the special issue. *Applied Cognitive Psychology*, 19, 233–241.
- Ernst, E. (2003). Researching the “absurd”. *Focus on Alternative and Complementary Therapies*, 8, 308–309.
- Ernst, E. (2008). The recent history of acupuncture. *The American Journal of Medicine*, 121, 1027–1028.
- Ernst, E., Lee, M. S., & Choi, T.-Y. (2011). Acupuncture for depression?: A systematic review of systematic reviews. *Evaluation and the Health Professions*, 34(4), 403–412.
- Errington-Evans, N. (2012). Acupuncture for anxiety. *CNS Neuroscience and Therapeutics*, 18, 277–284.
- Faigman, D. L., & Monahan, J. (2009). Standards of legal admissibility and their implications for psychological science. In J. L. Skeem, K. S. Douglas, & S. O. Lilienfeld (Eds.), *Psychological science in the courtroom: Consensus and controversy* (pp. 3–25). New York: Guilford.

- Fann, K. T. (1970). *Peirce's theory of abduction*. The Hague, Holland: Martinus Nijhoff.
- Faust, D. (2012). *Coping With Psychiatric and Psychological Testimony* (6th ed.). New York: Oxford University Press.
- Faust, D. (2007). Decision research can increase the accuracy of clinical judgment and thereby improve patient care. In S.O. Lilienfeld & W.T. O'Donohue (Eds.), *The great ideas of clinical science: 17 principles that every mental health professional should understand* (pp. 49–74). New York: Routledge.
- Faust, D., & Meehl, P. E. (2002). Using meta-scientific studies to clarify or resolve questions in the philosophy and history of science. *Philosophy of Science*, 69, S185–S196.
- Fava, J. L., & Velicer, W. F. (1992). The effects of overextraction on factor and component analysis. *Multivariate Behavioral Research*, 27, 387–415.
- Feinstein, D. (2008). Energy psychology: A review of the preliminary evidence. *Psychotherapy: Research, Practice, Training*, 45(2), 199–213.
- Feyerabend, P. (2002). *Against method* (3rd ed.). New York: Verso.
- Feynman, R. P. (1974). Cargo cult science: Some remarks on science, pseudoscience, and learning how not to fool yourself. Caltech's 1974 commencement address. *Engineering and Science*, 37(7), 10–13.
- Feynman, R. P. (1985). "Surely you're joking, Mr. Feynman:" *Adventures of a curious character*. New York: W.W. Norton.
- Feynman, R. P. (1998). *The meaning of it all: Thoughts of a citizen-scientist*. New York: Basic Books.

- Foa, E. B., Gillihan, S. J., & Bryant, R. A. (2013). Challenges and successes in dissemination of evidence-based treatments for posttraumatic stress: Lessons learned from prolonged exposure therapy for PTSD. *Psychological Science in the Public Interest*, 14(2), 65–111.
- Foa, E. B., & Meadows, E. A. (1997). Psychosocial treatments for posttraumatic stress disorder: A critical review. *Annual Review of Psychology*, 48, 449–480.
- Foa, E. B., & Zoellner, L. A., & Feeny, N. C. (2006). An evaluation of three brief programs for facilitating recovery after assault. *Journal of Traumatic Stress*, 19(1), 29–43.
- Fredrickson, R. (1992). *Repressed memories: A journey to recover from sexual abuse*. New York: Simon & Schuster.
- Gabel, J. (1976). *False consciousness: An essay on reification*. New York: Harper & Row.
- Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment*. Washington, DC: American Psychological Association.
- Garb, H. N., & Boyle, P. A. (2003). Understanding why some clinicians use pseudoscientific methods: Findings from research on clinical judgment. In S. O. Lilienfeld, S. J. Lynn, & J. M. Lohr (Eds.), *Science and pseudoscience in clinical psychology* (pp. 17–38). New York: Guilford Press.
- Gardner, M. (1957). *Fads and fallacies in the name of science: The curious theories of modern pseudoscientists and the strange, amusing and alarming cults that surround them. A study in human gullibility*. Mineola, NY: Dover.
- Garson, G. D. (2013). *Factor analysis*. Asheboro, NC: Statistical Associates.

- Gaudiano, B. A., Brown, L. A., & Miller, I. W. (2011). Factors associated with critical thinking abilities in psychotherapists. *Cognitive Behaviour Therapy, 40*, 137–146.
- Gaudiano, B. A., Brown, L. A., & Miller, I. W. (2012). Tapping their patients' problems away? Characteristics of psychotherapists using energy meridian techniques. *Research on Social Work Practice, 22*(6), 647–655.
- Gauvreau, P., & Bouchard, S. (2008). Preliminary evidence for the efficacy of EMDR in treating generalized anxiety disorder. *Journal of EMDR Practice and Research, 2*, 26–40.
- Gelso, C. J. (1985). Rigor, relevance, and counseling research: On the need to maintain our course between Scylla and Charybdis. *Journal of Counseling and Development, 63*, 551–554.
- Gist, R., & Woodall, S. J. (1995). Occupational stress in contemporary fire service. *Occupational Medicine: State of the Art Reviews, 10*, 763–787.
- Goldberg, L. R. (1965). Diagnosticians versus diagnostic signs: The diagnosis of psychosis versus neurosis from the MMPI. *Psychological Monographs, 79*(9), 602.
- Goldberg, L. R. (1968). Simple models or simple processes? Some research on clinical judgments. *American Psychologist, 23*, 483–496.
- Goldberg, L. R. (1970). Man versus model of man: A rationale, plus some evidence, for a method of improving on clinical inferences. *Psychological Bulletin, 73*, 422–432.
- Goldfried, M. R., Newman, M. G., Castonguay, L. G., Fuertes, J. N., Magnavita, J. J.,

- Sobell, L., & Wolf, A. W. (2014). On the dissemination of clinical experiences in using empirically supported treatments. *Behavior Therapy*, 45(1), 3–6.
- Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, 8, 206–213.
- Grant, M., & Threlfo, C. (2002). EMDR in the treatment of chronic pain. *Journal of Clinical Psychology*, 58, 1505–1520.
- Greasley, P. (2010). Is evaluating complementary and alternative medicine equivalent to evaluating the absurd? *Evaluation & the Health Professions*, 33(2), 127–139.
- Green, K. E., & Kvidahl, R. F. (1989). Personalization and offers of results: Effects on response rates. *The Journal of Experimental Education*, 57(3), 263–270.
- Grizzle, A. J., Olson, B. M., & Motheral, B. R. (2000). Therapeutic value: Who decides? *Pharmaceutical Executive*, 18, 84–90.
- 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(2), 293–323.
- Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12(1), 19–30.
- Grünbaum, A. (1985). Explication and implications of the placebo concept. In L.

- White, B. Tursky, & G. E. Schwartz (Eds.), *Placebo: Theory, research, and mechanisms* (pp. 9–36). New York: Guilford Press.
- Hand, W. D. (1985). Magical medicine: An alternative to “alternative medicine.” *Western Folklore*, 44, 240–251.
- Harcum, E. R., & Rosen, E. F. (1993). *The gatekeepers of psychology: Evaluation of peer review by case history*. Santa Barbara, CA: Praeger.
- Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (Eds.). (1997). *What if there were no significance tests?* Mahwah, NJ: Erlbaum.
- Hart, T., Fann, J. R., & Novack, T. A. (2008). The dilemma of the control condition in experience-based cognitive and behavioural treatment research. *Neuropsychological Rehabilitation*, 18, 1–21.
- Hastie, R., & Dawes, R. M. (2010). *Rational choice in an uncertain world: The psychology of judgment and decision making*. Thousand Oaks, CA: Sage.
- Hayes, S. C., Barlow, D. H., & Nelson-Gray, R. O. (1999). *The scientist practitioner: Research and accountability in the age of managed care*. Needham Heights, MA: Allyn & Bacon.
- Hays, K. A., Rardin, D. K., Jarvis, P. A., Taylor, N. M., Moorman, A. S., & Armstead, C. D. (2002). An exploratory survey on empirically supported treatments: Implications for internship training. *Professional Psychology: Research and Practice*, 33(2), 207–211.
- Heerwegh, D., Vanhove, T., Matthijs, K., & Loosveldt, G. (2005). The effect of personalization on response rates and data quality in Web surveys. *International Journal of Social Research Methodology*, 8(2), 85–99.

- Herbert, J. D., & Gaudiano, B. A. (2005). Moving from empirically supported treatment lists to practice guidelines in psychotherapy: The role of the placebo concept. *Journal of Clinical Psychology, 61*(7), 893–908.
- Herbert, J. D., Lilienfeld, S. O., Lohr, J. M., Montgomery, R. W., O'Donohue, W. T., Rosen, G. M., & Tolin, D. F. (2000). Science and pseudoscience in the development of eye movement desensitization and reprocessing: Implications for clinical psychology. *Clinical Psychology Review, 20*(8), 945–971.
- Hermann, R. C., Ettner, S. L., Dorwart, R. A., Langman-Dorwart, N., & Kleinman, S. (1999). Diagnoses of patients treated with ECT: A comparison of evidence-based standards with reported use. *Psychiatric Services, 50*(8), 1059–1065.
- Hershberger, S. L., & Fisher, D. G. (2003). A note on determining the number of imputations for missing data. *Structural Equation Modeling, 10*(4), 648–650.
- Hines, T. (2003). *Pseudoscience and the paranormal: A critical examination of the evidence*. Buffalo, NY: Prometheus.
- Hipol, L. J., & Deacon, B. J. (2012). Dissemination of evidence-based practices for anxiety disorders in Wyoming: A survey of practicing psychotherapists. *Behavior Modification, 37*(2), 170–188.
- Hodges, K., & Wotring, J. (2004). The role of monitoring outcomes in initiating implementation of evidence-based treatments at the state level. *Psychiatric Services, 55*(4), 396–400.
- Hoeppner, B. B., Redding, C. A., Rossi, J. S., Pallonen, U. E., Prochaska, J. O., & Velicer, W. F. (2012). Factor structure of Decisional Balance and Temptations Scales for Smoking: Cross-validation in urban female African-

- American adolescents. *International Journal of Behavioral Medicine*, 19, 217–227.
- Hofmann, S. G., & Lohr, J. M. (2010, January). To kill a Dodo bird. *The Behavior Therapist*, 33(1), 14–15.
- Hollis, J. F., Bills, R., Whitlock, E., Stevens, V. J., Mullooly, J., & Lichtenstein, E. (2000). Implementing tobacco interventions in the real world of managed care. *Tobacco Control*, 9(Suppl. 1), i18–i24.
- Holt, H., & Beutler, L. E. (2014). Concerns about the dissemination and implementation of evidence-based psychotherapies in the Veterans Affairs Health Care System. *American Psychologist*, 69(7), 705–706.
- Honyashiki, M., Furukawa, T. A., Noma, H., Tanaka, S., Chen, P., ...Caldwell, D. M. (2014). Specificity of CBT for depression: A contribution from multiple treatments meta-analyses. *Cognitive Therapy and Research*, 38(3), 249–260.
- Horan, W. P., & Blanchard, J. J. (2001, winter). Training opportunities in empirically supported treatments and their relationship to intern recruitment and post-internship placement: A survey of directors of internship training. *Clinical Science*, 1–10.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179–185.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Huber, P. W. (1993). *Galileo's revenge: Junk science in the courtroom*. New York:

BasicBooks.

- Hudson, J. I., Chase, E. A., & Pope, H. G. (1998). Eye movement desensitization and reprocessing in eating disorders: Caution against premature acceptance. *International Journal of Eating Disorders*, 23, 1–5.
- Hume, D. (1748/2007). Of Miracles. In Author, *An Enquiry Concerning Human Understanding* (pp. 99–118). Minneapolis: Filiquarian Publishing.
- Hunt, M. (1999). *The new know-nothings: The political foes of the scientific study of human nature*. New Brunswick, NJ: Transaction.
- Hutton, P., & Taylor, P. J. (2014). Cognitive behavioural therapy for psychosis prevention: A systematic review and meta-analysis. *Psychological Medicine*, 44(3), 449–468.
- Irwin, T. (1988). *Aristotle's first principles*. New York: Oxford University Press.
- Jaberghaderi, N., Greenwald, R., Rubin, A., Zand, S. O., & Dolatabadi, S. (2004). A comparison of CBT and EMDR for sexually-abused Iranian girls. *Clinical Psychology and Psychotherapy*, 11, 358–368.
- Jacobson, J. W., Mulick, J. A., & Schwartz, A. A. (1995). A history of facilitated communication: Science, pseudoscience, and antiscience science working group on facilitated communication. *American Psychologist*, 50(9), 750–765.
- Jones, J. L., & Mehr, S. L. (2007). Foundations and assumptions of the scientist-practitioner model. *American Behavioral Scientist*, 50(6), 766–771.
- Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). Thousand Oaks, CA: Sage.

- Jöreskog, K., & Sörbom, D. (1986). *LISREL VI: Analysis of linear structural relationships by maximum likelihood and least square methods*. Mooresville, IN: Scientific Software.
- Josefson, D. (2001). Rebirthing therapy banned after girl died in 70 minute struggle. *British Medical Journal*, 322(7293), 1014.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus, and Giroux.
- Karekla, M., Lundgren, J. D., & Forsyth, J. P. (2004). A survey of graduate training in empirically supported and manualized treatments: A preliminary report. *Cognitive and Behavioral Practice*, 11, 230–242.
- Kaslow, N. J., Borden, K. A., Collins, F. L., Jr., Forrest, L., Illfelder-Kaye, J., Nelson, P., & Rallo, J. S. (2004). Competencies conference: Future directions in education and credentialing in professional psychology. *Journal of Clinical Psychology*, 60, 699–712.
- Kauth, M. R., Sullivan, G., Cully, J., & Blevins, D. (2011). Facilitating practice changes in mental health clinics: A guide for implementation development in health care systems. *Psychological Services*, 8(1), 36–47.
- Kazdin, A. E. (2003). *Research design in clinical psychology* (4th ed.). Boston, MA: Allyn and Bacon.
- Keane, T. M. (1998). Psychological and behavioral treatment of posttraumatic stress disorder. In P. Nathan and J. Gorman (Eds.), *A guide to treatments that work* (pp. 398–407). Oxford: Oxford University Press.
- Keane, T. M., & Barlow, D. H. (2002). Posttraumatic stress disorder. In: D.H.

- Barlow (Ed.), *Anxiety and its disorders: The nature and treatment of anxiety and panic* (2nd ed.). New York: Guilford Press.
- Keenan, P. S., Farrell, D. P. (2000). Treating morbid jealousy with eye movement desensitization and reprocessing utilizing cognitive inter-weave: A case report. *Counseling Psychology Quarterly*, 13, 175–189.
- Klerman, G. L. (1990). The psychiatric patient's right to effective treatment: Implications of *Osheroff v. Chestnut Lodge*. *Journal of Psychiatry*, 147, 409–418.
- Kline, M. (1967). *Mathematics for the nonmathematician*. New York: Dover Publications.
- Kohout, J., & Wicherski, M. (2010, October). 2011 graduate study in psychology snapshot: Applications, acceptances, enrollments, and degrees awarded to Master's- and doctoral-level students in U.S. and Canadian graduate departments of psychology: 2009–2010. Center for Workforce Studies: American Psychological Association. Retrieved from <http://www.apa.org/workforce/publications/11-grad-study/applications.pdf>
- Koocher, G. P., & Keith-Spiegel, P. (2008). Psychotherapy part II: Techniques and controversies. In Authors, *Ethics in Psychology and the Mental Health Professions: Standards and Cases* (3rd ed., pp. 127–158). New York: Oxford University Press.
- Krigolson, O. E., Pierce, L. J., Holroyd, C. B., & Tanaka, J. W. (2008). Learning to become an expert: Reinforcement learning and the acquisition of perceptual expertise. *Journal of Cognitive Neuroscience*, 21(9), 1834–1841.

- Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–196). London: Cambridge University Press.
- Lakatos, I. (1974). Science and pseudoscience. *Conceptus: Zeitschrift Fuer Philosophie*, 8, 5–9.
- Lambert, M. J. (Ed.). (2013). *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.). Hoboken, NJ: Wiley.
- Laudan, L. (1996). *Beyond positivism and relativism: Theory, method, and evidence*. Boulder, CO: Westview Press.
- Laudan, L. (1983). The demise of the demarcation problem. In R. S. Cohen & L. Laudan (Eds.), *Physics, philosophy and psychoanalysis: Essays in honor of Adolf Grünbaum* (pp. 111–127). Dordrecht, Netherlands: Reidel.
- Laurenceau, J-P., Hayes, A. M., & Feldman, G. C. (2007). Some methodological and statistical issues in the study of change processes in psychotherapy. *Clinical Psychology Review*, 27, 682–695.
- Leahey, T. H., & Leahey, G. E. (1983). *Psychology's occult doubles: Psychology and the problem of pseudoscience*. Chicago: Nelson-Hall.
- Lewandowsky, S., Little, D., & Kalish, M. L. (2007). Knowledge and expertise. In F. T. Durso (Ed.), *Handbook of applied cognition* (2nd ed., pp. 83–109). West Sussex, England: Wiley.
- Lewin, A. B., Wu, M. S., McGuire, J. F., & Storch, E. A. (2014). Cognitive behavior therapy for obsessive-compulsive and related disorders. *Psychiatric Clinics of North America*. Advance online publication.

- Lilienfeld, S. O. (2010). Can psychology become a science? *Personality and Individual Differences*, 49(4), 281–288.
- Lilienfeld, S. O. (2013, March). Closing the science-practice gap: Reflections from the armchair. *Observer* 26(3). Retrieved from <http://www.psychologicalscience.org/index.php/publications/observer/2013/march-13/closing-the-science-practice-gap.html>
- Lilienfeld, S. O. (2008). EMDR treatment: Less than meets the eye? In: S.O. Lilienfeld, J. Ruscio, and S.J. Lynn (Eds.), *Navigating the mindfield: A guide to separating science from pseudoscience in mental health* (pp. 403–418). New York: Prometheus Books.
- Lilienfeld, S. O. (1998, fall). Pseudoscience in contemporary clinical psychology: What it is and what we can do about it. *The Clinical Psychologist*, 51(4), 3–9.
- Lilienfeld, S. O. (1997). Psychological treatments that cause harm. *Perspectives on Psychological Science*, 2(1), 53–70.
- Lilienfeld, S. O. (2012). Public skepticism of psychology: Why many people perceive the study of human behavior as unscientific. *American Psychologist*, 67(2), 111–129.
- Lilienfeld, S. O. (2005). The 10 commandments of helping students distinguish science from pseudoscience in psychology. *Observer*, 18(9), 39–40 and 49–51.
- Lilienfeld, S. O., Lynn, S. J., & Lohr, J. M. (2003). *Science and pseudoscience in clinical psychology*. New York: Guilford Press.
- Lilienfeld, S. O., Lynn, S. J., Ruscio, J., & Beyerstein, B. L. (2010). *50 great myths*

of popular psychology: Shattering widespread misconceptions about human behavior. Malden, MA: Wiley.

Lilienfeld, S. O., Ruscio, J., & Lynn, S. J. (2008). *Navigating the mindfield: A guide to separating science from pseudoscience in mental health.* New York: Prometheus Books.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York: Wiley.

Litz, B. T., Gray, M. J., Bryant, R. A., & Adler, A. B. (2002). Early intervention for trauma: Current status and future directions. *Clinical Psychology: Science and Practice*, 9, 112–134.

Loehlin, J. C. (2004). *Latent variable models: An introduction to factor, path, and structural equation analysis* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Loftus, E. (1995). Remembering dangerously. *Skeptical Inquirer*, 19(2), 20–29.

Loftus, E. F. (1993). The reality of repressed memories. *American Psychologist*, 48, 518–537.

Loftus, E., & Ketcham, K. (1994). *The myth of repressed memory: False memories and allegations of sexual abuse.* New York: St. Martin's Press.

Lohr, J. M., Fowler, K. A., & Lilienfeld, S. O. (2002, summer). The dissemination and promotion of pseudoscience in clinical psychology: The challenge to legitimate clinical science. *The Clinical Psychologist*, 55(3), 4–10.

Lohr, J. M., Hooke, W., Gist, R., & Tolin, D. F. (2003). Novel and controversial treatments for trauma-related stress disorders. In: Lilienfeld, S.O., Lynn, S.J.,

- and Lohr, J.M. (Eds.), *Science and pseudoscience in clinical psychology* (pp. 243–272). New York: Guilford Press.
- Lohr, J. M., Olatunji, B. O., Parker, L., & DeMaio, C. (2005). Experimental analysis of specific treatment factors: Efficacy and practice implications. *Journal of Clinical Psychology, 61*(7), 819–834.
- Lohr, J. M., Tolin, D. F., & Kleinknecht, R. A. (1996). An intensive design investigation of eye movement desensitization and reprocessing of claustrophobia. *Journal of Anxiety Disorders, 10*, 73–88.
- Lohr, J. M., Tolin, D. F., & Lilienfeld, S. O. (1998). Efficacy of eye movement desensitization and reprocessing: Implications for behavior therapy. *Behavior Therapy, 29*, 123–156.
- Lowman, R. L. (2012). The scientist-practitioner consulting psychologist. *Consulting Psychology Journal: Practice and Research, 64*(3), 151–156.
- Lynn, S. J., Lock, T., Loftus, E. F., Krackow, E., & Lilienfeld, S. O. (2003). The remembrance of things past: Problematic memory recovery techniques in psychotherapy. In S. O. Lilienfeld, S. J. Lynn, & J. M. Lohr (Eds.), *Science and pseudoscience in clinical psychology* (pp. 205–239). New York: Guilford Press.
- Lynn, S. J., Loftus, E. F., Lilienfeld, S. O., & Lock, T. (2008). Memory recovery techniques in psychotherapy: Problems and pitfalls. In S. O. Lilienfeld, J. Ruscio, & S. J. Lynn (Eds.), *Navigating the mindfield: A guide to separating science from pseudoscience in mental health* (pp. 283–302). New York: Prometheus Books.

- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130–149.
- Maciosek, M. V., Coffield, A. B., Edwards, N. M., Flottemesch, T. J., Goodman, M. J., ...Solberg, L. I. (2006). Priorities among effective clinical preventive services: Results of a systematic review and analysis. *American Journal of Preventative Medicine, 31*(1), 52–61.
- Majima, Y. (2012). The role of critical thinking in acquisition of science literacy and “subjective demarcation” between science and pseudoscience. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society, 19*(1), 22–38.
- Maran, M. (2010). *My lie: The true story of false memory*. San Francisco, CA: Wiley.
- Marcus, D., & McCullough, L. (2009). An evaluation of the evidence in “evidence-based” integrative medicine programs. *Academic Medicine, 84*(9), 1229–1234.
- Marich, J. (2010). Eye movement desensitization and reprocessing in addiction continuing care: A phenomenological study of women in recovery. *Psychology of Addictive Behaviors, 24*(3), 498–507.
- Marino, L., & Lilienfeld, S. O. (2007). Dolphin-assisted therapy: More flawed data and more flawed conclusions. *Anthrozoös, 20*(3), 239–249.
- Marsa, L. (2002, March 25). Trauma therapy’s new focus. *Los Angeles Times*. Retrieved from <http://articles.latimes.com/2002/mar/25/health/he-emdr25>
- Marten, P. A., & Heimberg, R. G. (1995). Toward an integration of independent practice and clinical research. *Professional Psychology: Research and*

Practice, 26(1), 48–53.

Mayou, R. A., Ehlers, A., & Hobbs, M. (2000). Psychological debriefing for road traffic accident victims. *British Journal of Psychiatry*, 176, 589–593.

McLean, C. P., & Miller, N. A. (2010). Changes in critical thinking skills following a course on science and pseudoscience: A quasi-experimental study. *Teaching of Psychology*, 37, 85–90.

McFall, R. M. (1991). Manifesto for a science of clinical psychology. *The Clinical Psychologist*, 44, 75–88.

McHugh, R. K., & Barlow, D. H. (2010). The dissemination and implementation of evidence-based psychological treatments: A review of current efforts. *American Psychologist*, 65(2), 73–84.

McIlvried, E. J., Wall, J. R., Kohout, J., Keys, S., & Goreczny, A. (2010). Graduate training in clinical psychology: Student perspectives on selecting a program. *Training and Education in Professional Psychology*, 4(2), 105–115.

McKenzie, D. (1986). Teaching students who already know the “truth.” *Thought & Action*, 2, 109–122.

McNally, R. J. (1999). EMDR and Mesmerism: A comparative historical analysis. *Journal of Anxiety Disorders*, 13(1–2), 225–236.

McNally, R. J. (2003). Is the pseudoscience concept useful for clinical psychology? The demise of pseudoscience. *The Scientific Review of Mental Health Practice*, 2(2), 97–101.

McNally, R. J. (2004). Is traumatic amnesia nothing but psychiatric folklore? *Cognitive Behavior Therapy*, 33, 97–101.

- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.
- Meehl, P. E. (1993). Philosophy of science: Help or hindrance? *Psychological Reports*, 72, 707–733.
- 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, 806–814.
- Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports*, 66, 195–244.
- Meehl, P. E., & MacCorquodale, K. (1991). On a distinction between hypothetical constructs and intervening variables. In C.A. Anderson & K. Gunderson (Eds.), *Paul E. Meehl: Selected philosophical and methodological papers* (pp. 249–264). Minneapolis, MN: University of Minnesota Press.
- Mercer, J. (2008). Attachment therapy: A treatment without empirical support. In S. O. Lilienfeld, J. Ruscio, & S. J. Lynn (Eds.), *Navigating the mindfield: A guide to separating science from pseudoscience in mental health* (pp. 435–453). New York: Prometheus Books.
- Meyer, J. F., Baker, A. M., & Baker, N. J. (2012, March). *Re-goring the sacred cow of clinical judgment: Why such resistance, and what can we do about it?* Interactive poster session presented at the third biennial midwinter conference for Division 24 of the American Psychological Association: Society for Theoretical and Philosophical Psychology, Austin, TX.
- McFall, R. M. (1991). Manifesto for a science of clinical psychology. *The Clinical*

- Psychologist*, 44(6), 75–88.
- Michalski, D. S., & Kohout, J. L. (2011). The state of the psychology health service provider workforce. *American Psychologist*, 66(9), 825–834.
- Miranda, J., Nakamura, R., & Bernal, G. (2003). Including ethnic minorities in mental health intervention research: A practical approach to a long-standing problem. *Culture, Medicine, and Psychiatry*, 27, 467–486.
- Mischel, W. (2009). Connecting clinical practice to scientific progress. *Psychological Science in the Public Interest*, 9(2), i–ii.
- Mitchell, K. R. (1977). Professional training in psychology: Some history, themes, and comment. *Australian Psychologist*, 12(1), 89–94.
- Morier, D., & Keeports, D. (1994). Normal science and the paranormal: The effect of a scientific method course on students' beliefs. *Research in Higher Education*, 35(4), 443–453.
- Morrison, D. E., & Henkel, R. E. (Eds.). (2009). *The significance test controversy*. New Brunswick, NJ: Transaction.
- Muris, P., Merkelbach, H., van Haaften, H., & Mayer, B. (1997). Eye movement desensitization and reprocessing versus exposure in vivo: A single-session crossover study of spider-phobic children. *British Journal of Psychiatry*, 171, 82–86.
- Muthén, L. K., & Muthén, B. O. (1998-2009). *Mplus 5.0 (Computer software)*. Los Angeles: Author.
- Nathan, D., & Snedeker, M. (2001). *Satan's silence: Ritual abuse and the making of a modern American witch hunt*. Lincoln, NE: iUniverse.com

- Nathan, P. E., & Gorman, J. M. (Eds.). (2007). *A guide to treatments that work* (3rd ed.). New York: Oxford University Press.
- Neimeyer, R. (2000). Searching for the meaning of meaning: Grief therapy and the process of reconstruction. *Death Studies*, 24, 541–558.
- Norcross, J. C. (Ed.). (2011). *Psychotherapy relationships that work: Evidence-based responsiveness* (2nd ed.). New York: Oxford University Press.
- Norcross, J. C., Hanych, J. M., & Terranova, R. D. (1996). Graduate study in psychology: 1992–1993. *American Psychologist*, 51(6), 631–643.
- O’Connell, N. E., Wand, B. M., & Goldacre, B. (2009). Interpretive bias in acupuncture research? A case study. *Evaluation & the Health Professions*, 32, 393–409.
- O’Connor, D. J. (1975). *The correspondence theory of truth*. London, UK: Hutchinson.
- O’Connor, K., Stip, E., Pélissier, M-C., Aardema, F., Guay, S., Gaudette, G., ...Leblanc, V. (2007). Treating delusional disorder: A comparison of cognitive-behavioural therapy and attention placebo control. *The Canadian Journal of Psychiatry*, 52(3), 182–190.
- Olatunji, B. O., Kauffman, B. Y., Meltzer, S., Davis, M. L., Smits, J. A. J., & Powers, M. B. (2014). Cognitive-behavioral therapy for hypochondriasis/health anxiety: A meta-analysis of treatment outcome and moderators. *Behaviour Research and Therapy*, 58, 65–74.
- Olatunji, B. O., Parker, L. M., & Lohr, J. M. (2005–2006, fall/winter). Pseudoscience in contemporary psychology: Professional issues and implications. *The*

Scientific Review of Mental Health Practice, 4(2), 19–36.

Olinsky, A., Chen, S., & Harlow, L. (2003). The comparative efficacy of imputation methods for missing data in structural equation modeling. *European Journal of Operational Research*, 151, 53–79.

Oltmanns, T. F., & Klonsky, E. D. (2007). Critical thinking in clinical inference. In R. J. Sternberg, H. L. Roediger III, & D. F. Halpern (Eds.), *Critical thinking in psychology* (pp. 196–215). New York: Cambridge University Press.

Omer, H., & London, P. (1989). Signal and noise in psychotherapy: The role and control of non-specific factors. *The British Journal of Psychiatry*, 155, 239–245.

Orr, L. (1998). *Breaking the death habit: The science of everlasting life*. Berkeley, CA: Frog Books.

Otto, M. W., & Deveney, C. (2005). Cognitive-behavioral therapy and the treatment of panic disorder: Efficacy and strategies. *Journal of Clinical Psychiatry*, 66(Suppl. 4), 28–32.

Overholser, J. C. (2010). Clinical expertise: A preliminary attempt to clarify its core elements. *Journal of Contemporary Psychotherapy*, 40(3), 131–139.

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, 76(6), 972–987.

Peirce, C. S. (2011). The scientific attitude and fallibilism. In J. Buchler (Ed.), *Philosophical writings of Peirce* (pp. 42–59). Mineola, NY: Dover.

Perkins, B. R., & Rouanzoin, C. C. (2002). A critical evaluation of current views

- regarding Eye Movement Desensitization and Reprocessing (EMDR):
Clarifying points of confusion. *Journal of Clinical Psychology*, 58, 77–97.
- Peterson, D. R. (2000). Scientist-practitioner or scientific practitioner? *American Psychologist*, 55(2), 252–253.
- Peterson, D. R. (2005). *Twelve years of correspondence with Paul Meehl: Tough notes from a gentle genius*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Pigliucci, M., & Boudry, M. (Eds.). (2013). *Philosophy of pseudoscience: Reconsidering the demarcation problem*. Chicago, IL: University of Chicago Press.
- Pignotti, M., & Thyer, B. (2009). Some comments on “Energy psychology: A review of the evidence”: Premature conclusions based on incomplete evidence? *Psychotherapy Theory, Research, Practice, Training*, 46, 257–261.
- Pinsof, W. M., Goldsmith, J. Z., & Latta, T. A. (2012). Information technology and feedback research can bridge the scientist-practitioner gap: A couple therapy example. *Couple and Family Psychology: Research and Practice*, 1(4), 253–273.
- Poissal, J. A., Truffer, C., Smith, S., Sisko, A., Cowan, C., Keehan, S. ...the National Health Expenditure Accounts Projections Team. (2007). Health spending projections through 2016: Modest changes obscure part D’s impact. *Health Affairs*, 26(2), w242–w253.
- Popper, K. A. (2002). *Conjectures and refutations*. New York: Routledge Classics.
- Posmontier, B., Dovydaitis, T., & Lipman, K. (2010). Sexual violence: Psychiatric healing with eye movement reprocessing and desensitization. *Health Care for*

- Women International*, 31, 755–768.
- Pratkanis, A. R. (1995, July/August). How to sell a pseudoscience. *Skeptical Inquirer*, 19(4), 19–25.
- Protinsky, H., Flemke, K., & Sparks, J. (2001). EMDR and emotionally oriented couples therapy. *Contemporary Family Therapy*, 23, 153–168.
- Psychology: A reality check. (2009). *Nature*, 461(7266).
- Quinlan, K. B., & McCaul, K. D. (2000). Matched and mismatched interventions with young adult smokers: Testing a stage theory. *Health Psychology*, 19(2), 165–171.
- Raimy, V. C. (1950). *Training in clinical psychology*. New York: Prentice Hall.
- Renfrey, G., & Spates, C. R. (1994). Eye movement desensitization: A partial dismantling study. *Journal of Behavior Therapy and Experimental Psychiatry*, 25, 231–239.
- Resnik, D. B. (2000). A pragmatic approach to the demarcation problem. *Studies in History and Philosophy of Science*, 31(2), 249–267.
- Ricci, R. J., Clayton, C. A., & Shapiro, F. (2006). Some effects of EMDR on previously abused child molesters: Theoretical reviews and preliminary findings. *The Journal of Forensic Psychiatry and Psychology*, 17, 538–562.
- Rodenburg, R., Benjamin, A., Meijer, A. M., & Jongeneel, R. (2009). Eye movement desensitization and reprocessing in an adolescent with epilepsy and mild intellectual disability. *Epilepsy and Behavior*, 16, 175–180.
- Roth, A., & Fonagy, P. (2005). *What works for whom? A critical review of psychotherapy research* (2nd ed.). New York: Guilford.

- Rosch, E. (1973). Natural categories. *Cognitive Psychology*, 4, 328–350.
- Rothbaum, B. O., Astin, M. C., & Marsteller, F. (2005). Prolonged exposure versus eye movement desensitization and reprocessing (EMDR) for PTSD rape victims. *Journal of Traumatic Stress*, 18, 607–616.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63, 581–592.
- Ruscio, J. (2005). *Critical thinking in psychology: Separating sense from nonsense* (2nd ed.). Belmont, CA: Wadsworth.
- Russell, B. (1912). *The problems of philosophy*. Cambridge, England: Home University Library.
- Sagan, C. (1995, January/February). Wonder and skepticism. *Skeptical Inquirer*, 19.1. Retrieved from http://www.csicop.org/si/show/wonder_and_skepticism
- Sagan, C., & Druyan, A. (1997). *The demon-haunted world: Science as a candle in the dark*. New York: Random House.
- Samuels, A., & Spradlin, T. (1995). Quantitative behavioral study of bottlenose dolphins in swim-with-the-dolphin programs in the United States. *Marine Mammal Science*, 11, 520–544.
- Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological Bulletin*, 66, 178–200.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.
- Schurz, G. (2014). *Philosophy of science: A unified approach*. New York: Routledge.
- Seaton, B. E., Allen, D. N., Goldstein, G., Kelley, M. E., & van Kammen, D. P.

- (1999). Relations between cognitive and symptom profile heterogeneity in schizophrenia. *The Journal of Nervous and Mental Disease*, 187(7), 414–419.
- Seery, M. D. (2011). Resilience: A silver lining to experiencing adverse life events? *Current Directions in Psychological Science*, 20(6), 390–394.
- Shackel, N. (2013). Pseudoscience and idiosyncratic theories of rational belief. In M. Pigliucci & M. Boudry (Eds.), *Philosophy of pseudoscience: Reconsidering the demarcation problem* (pp. 417–438). Chicago, IL: University of Chicago Press.
- Shakow, D. (1969). Homo scientius et homo professionalis—sempervirens? In D. Shakow (Ed.), *Clinical psychology as science and profession: A forty-year odyssey* (pp. 176–184). Chicago: Aldine.
- Shakow, D. (1976). What is clinical psychology? *American Psychologist*, 31, 553–560.
- Shapiro, F. (1989a). Eye movement desensitization: A new treatment for post-traumatic stress disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, 20, 211–217.
- Shapiro, F. (1989b). Efficacy of the eye-movement desensitization procedure in the treatment of traumatic memories. *Journal of Traumatic Stress*, 2, 199–223.
- Shapiro, F. (1991). Eye movement desensitization and reprocessing procedure: From EMD to EMD/R—A new treatment model for anxiety and related traumata. *Behavior Therapist*, 14, 133–135.
- Shapiro, F. (1994a). Alternative stimuli in the use of EMD(R). *Journal of Behaviour Therapy and Experimental Psychiatry*, 25, 89–91.

- Shapiro, F. (1994b). EMDR: In the eye of a paradigm shift. *Behavior Therapist*, 17, 153–156.
- Shapiro, F. (1995). *Eye movement desensitization and reprocessing: Basic principles, protocols, and procedures*. New York: Guilford Press.
- Shapiro, F. (1999). Eye movement desensitization and reprocessing (EMDR) and the anxiety disorders: Clinical and research implications of an integrated psychotherapy treatment. *Journal of Anxiety Disorders*, 13, 35–67.
- Shapiro, F., & Forrest, M. S. (2004). *EMDR: The breakthrough therapy for overcoming anxiety, stress, and trauma*. New York: Basic Books.
- Sharp, I. R., Herbert, J. D., & Redding, R. E. (2008). The role of critical thinking skills in practicing clinical psychologists' choice of intervention techniques. *The Scientific Review of Mental Health Practice: Objective Investigations of Controversial and Unorthodox Claims in Clinical Psychology, Psychiatry, and Social Work*, 6(1), 21–30.
- Sheaffer, R. (2008). The fallacy of misplaced rationalism. *Skeptical Inquirer*, 32, 23–24, 47.
- Sheffer, C. E., Stitzer, M., Payne, T. J., Applegate, B. W., Bourne, D., & Wheeler, J. G. (2009). Treatment for tobacco dependence for rural, lower-income smokers: Outcomes, predictors and measurement considerations. *American Journal of Health Promotion*, 23(5), 328–338.
- Shermer, M. (2002). *Why people believe weird things: Pseudoscience, superstition, and other confusions of our time*. New York: Henry Holt and Company.
- Shiffman, S., Brockwell, S. E., Pillitteri, J. L., & Gitchell, J. G. (2008). Use of

- smoking-cessation treatments in the U.S. *American Journal of Preventative Medicine*, 34(2), 102–111.
- Singer, M., & Lalich, J. (2008). Back to the beginning: Regression, reparenting, and rebirthing. In S. O. Lilienfeld, J. Ruscio, & S. J. Lynn (Eds.), *Navigating the mindfield: A guide to separating science from pseudoscience in mental health* (pp. 455–477). New York: Prometheus Books.
- Singer, M. T., & Lalich, J. (1996). “Crazy” therapies: What are they? Do they work? San Francisco, CA: Jossey-Bass.
- Singer, M. T., & Nievod, A. (2003). New Age therapies. In S. O. Lilienfeld, S. J. Lynn, & J. M. Lohr (Eds.), *Science and pseudoscience in clinical psychology* (pp. 176–204). New York: Guilford Press.
- Slack, R. (2010, June). Acupuncture: A science-based assessment. A position paper from the Center for Inquiry Office of Public Policy. Retrieved from http://www.csicop.org/specialarticles/show/acupuncture_a_science-based_assessment/
- Sonne-Holm, S., Sørensen, T. I. A., Jensen, G., & Schnohr, P. (1989). Influence of fatness, intelligence, education and sociodemographic factors on response rate in a health survey. *Journal of Epidemiology and Community Health*, 43, 369–374.
- Stanovich, K. E. (2003). *How to think straight about psychology* (7th ed.). New York, NY: HarperCollins.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173–180.

- Stewart, R. E., Chambless, D. L., & Baron, J. (2012). Theoretical and practical barriers to practitioners' willingness to seek training in empirically supported treatments. *Journal of Clinical Psychology*, 68(1), 8–23.
- Still, A., & Dryden, W. (2004). The social psychology of “pseudoscience”: A brief history. *Journal for the Theory of Social Behaviour*, 34(3), 265–290.
- Swartz, S. H., Cowan, T. M., Klayman, J. E., Welton, M. T., & Leonard, B. A. (2005). Use and effectiveness of tobacco telephone counseling and nicotine therapy in Maine. *American Journal of Preventative Medicine*, 29(4), 288–294.
- Task Force on the Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically-validated treatments. *The Clinical Psychologist*, 48, 3–23.
- Tavris, C. (2003). The widening scientist-practitioner gap: A view from the bridge. In S. O. Lilienfeld, S. J. Lynn, & J. M. Lohr (Eds.), *Science and pseudoscience in clinical psychology* (pp. ix–xvii). New York: Guilford Press.
- Teachman, B. A., Drabick, D. A. G., Hershenberg, R., Vivian, D., Wolfe, B. E., & Goldfriend, M. R. (2012). Bridging the gap between clinical research and clinical practice: Introduction to the special edition. *Psychotherapy*, 49(2), 97–100.
- ter Laak, J., de Goede, M., Aleva, A., & van Rijswijk, P. (2005). The Draw-A-Person Test: An indicator of children's cognitive and socioemotional adaptation. *The Journal of Genetic Psychology*, 166(1), 77–93.
- Thagard, P. R. (1993). *Computational philosophy of science*. Cambridge, MA: MIT

Press.

- Thagard, P. R. (1980). Resemblance, correlation and pseudo-science. In M. P. Hanen, M. J. Osler, & R. G. Weyant (Eds.), *Science, pseudo-science and society* (pp. 17–28). Waterloo, Ontario: Wilfred Laurier University Press.
- Thagard, P. R. (1978). Why astrology is a pseudoscience. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 1*, 223–234.
- Tinker, R. H., & Wilson, S. A. (1999). *Through the eyes of a child: EMDR with children*. New York: Norton.
- Tolin, D. (2013, May 28). *Science, pseudoscience, and mental health treatment* [PowerPoint presentation]. Stockholm Psychiatry Lecture. Retrieved from https://www.youtube.com/watch?v=trt_tsTW8WY
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1–10.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.
- Unschuld, P. U. (2003). *Huang di nei jing su wen: Nature, knowledge, imagery in an ancient Chinese medical text*. Los Angeles: University of California Press.
- VanderVeen, J. W., Reddy, L. F., Veilleux, J. C., January, A. M., & DiLillo, D. (2012). Clinical PhD graduate student views of their scientist practitioner training. *Journal of Clinical Psychology*, 68(9), 1048–1057.
- Velicer, W. F., & Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, 3(2), 231–251.
- Victor, J. S. (1993). *Satanic panic: The creation of a contemporary legend*. Chicago:

Open Court.

- Vollmer, S., Spada, H., Caspar, F., & Burn, S. (2013). Expertise in clinical psychology: The effects of university training and practical experience on expertise in clinical psychology. *Frontiers in Psychology*, 4, 1–12.
- Vonnegut, K. (1973). *Wampeters, foma, and granfalloon (opinions)*. New York: Dell Publishing.
- Vyse, S. A. (1997). *Believing in magic: The psychology of superstition*. New York: Oxford University Press.
- Wachtel, P. L. (2010). Beyond “ESTs”: Problematic assumptions in the pursuit of evidence-based practice. *Psychoanalytic Psychology*, 27(3), 251–272.
- Wakefield, H., & Underwager, R. (1992). Recovered memories of alleged sexual abuse: Lawsuits against parents. *Behavioral Sciences and the Law*, 10, 483–507.
- Walters, B. (1995). When all else fails: New therapy using eye movements [Television broadcast]. In D. Sloan [Executive producer], *20/20*. New York: ABC News.
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ: Lawrence Erlbaum.
- Wampold, B. E., Imel, Z. E., Laska, K. M., Benish, S., Miller, S. D., Flückiger, C., ...Budge, S. (2010). Determining what works in the treatment of PTSD. *Clinical Psychology Review*, 30(8), 923–933.
- Watson, G., & Glaser, E. M. (1994). *Watson-Glaser Critical Thinking Appraisal-Manual Form S*. San Antonio, TX: Psychological Corporation.

- Weck, F., Weigel, M., Richtberg, S., & Stangier, U. (2011). Reliability of adherence and competence assessment in psychoeducational treatment: Influence of clinical experience. *Journal of Nervous and Mental Disease, 199*(12), 983–986.
- Wesp, R., & Montgomery, K. (1998). Developing critical thinking through the study of paranormal phenomena. *Teaching of Psychology, 4*, 275–278.
- Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin, 130*, 631–663.
- Witkowski, T. (2012). A review of research findings on neuro-linguistic programming. *The Scientific Review of Mental Health Practice: Objective Investigations of Controversial and Unorthodox Claims in Clinical Psychology, Psychiatry, and Social Work, 9*(1), 29–40.
- Wood, J. M., Nezworski, T., Lilienfeld, S. O., & Garb, H. N. (2003). *What's wrong with the Rorschach? Science confronts the controversial inkblot test*. San Francisco, CA: Wiley.
- World Health Organization. (2013). *Guidelines for the management of conditions specifically related to stress*. Retrieved from http://apps.who.int/iris/bitstream/10665/85119/1/9789241505406_eng.pdf
- World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *The Journal of the American Medical Association, 310*(20), 2191–2194.
- Worrall, R. S. (1990). Detecting health fraud in the field of learning disabilities.

Journal of Learning Disabilities, 23(4), 207–212.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34, 806–838.

Young, F. E. (1988, March). Allies in the war against health fraud. *The FDA Consumer*, 6–7.

Yu, C. H., Popp, S. O., DiGangi, S., & Jannasch-Pennell, A. (2007). Assessing unidimensionality: A comparison of Rasch modeling, parallel analysis, and TETRAD. *Practical Assessment, Research & Evaluation*, 14(12),

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99, 432–442.