In press at *Professional Psychology: Research and Practice*

Why Psychologists Should Reject Complementary and Alternative Medicine:

A Science-Based Perspective

Lawton K. Swan, Sondre Skarsten, and Martin Heesacker

University of Florida

John R. Chambers

Saint Louis University

Correspondence concerning this article should be addressed to Lawton K. Swan, 945 Center Drive, P.O. Box 112250, Gainesville, FL, 32611. Email: lkswan@ufl.edu

Abstract

Professional psychology is in apparent conflict about its relationship to "complementary" and "alternative" medicine (CAM)—some scholars envision a harmonious partnership, whereas others perceive irreconcilable differences. We propose that the field's ambivalence stems at least partly from the fact that inquiring psychologists can readily point to peer-reviewed empirical evidence (e.g., published reports of randomized controlled trials) to either substantiate or refute claims for the efficacy of most CAM modalities. Thankfully, recent intellectual developments in the fields of medicine and scientific psychology—developments which we refer to collectively as the *science-based perspective*—have led to the identification of several principles that may be used to judge the relative validity of conflicting health intervention research findings, including the need to consider (a) the prior scientific plausibility of a treatment's putative mechanism-of-action; and, commensurately, (b) the degree of equivalence between treatment and control groups—except for the single active element of the treatment believed to cause a specific change, all else between the two groups should be identical. To illustrate the potential of this approach to resolve psychology's CAM controversy, we conducted a re-review of the research cited by Barnett and Shale [2012; *Professional Psychology: Research and Practice, 43(6)*, 576-585] regarding the efficacy of 11 types of CAM that psychologists might endorse. Less than 15% of the studies we reviewed (*N* = 239) employed research designs capable of ruling out non-specific effects, and those that did tended to produce negative results. From a science-based perspective, psychologists should reject CAM in principle and practice.

Keywords*:* complementary, alternative, medicine, practice, ethics

Why Psychologists Should Reject Complementary and Alternative Medicine:

A Science-Based Perspective

How should professional psychologists regard the notion of “complementary” and “alternative” medicine (CAM)? According to one high-profile analysis of the relevant issues (Hughes, 2008), CAM *by definition* falls outside of psychology’s epistemic jurisdiction, insofar as psychologists identify themselves as *evidence-based* healthcare providers (APA, 2006). Namely, Hughes argued that CAM’s many products and practices—bound together nominally by their absence in the widely accepted canon of verified methods for preventing or curing disease—(a) have failed to demonstrate efficacy in high-quality research, and (b) very often rely on empirically untenable assumptions about basic physiological mechanisms and processes (e.g., acupuncture’s concept of *qì*, “the vital energy of every living organism and the source of all movement and change in the universe”; Dale, 1994). Given the considerable amount of effort that psychology has devoted to ensuring that its professionals base their clinical interventions in the philosophy and methods of natural science (Baker, McFall, & Shoham, 2009), the two fields would seem to be irreconcilably at odds.

Yet, calls for the integration of CAM into psychological practice have not ceased in the time since Hughes’ unfavorable treatise (Barnett & Shale, 2012; 2013; Barnett, Shale, Elkins, & Fisher, 2014; Hymel & Rich, 2014; Park, 2013). Barnett and Shale (2012), for instance, apparently found sufficient cause to recommend a broad array of CAM modalities to *Professional Psychology: Research and Practice* readers as promising new clinical competencies. In some cases, Barnett and Shale even encouraged practitioners to seek credentialing so that they might purvey CAM to their clients directly. Similar arguments for a marriage between professional psychology and CAM date back to at least the early 2000s (e.g., Bassman & Uellendahl, 2003; White, 2000), each clearly expressing confidence that *CAM can in fact comport with psychology's commitment to practice guided by objective, potentially falsifiable evidence*[1]. That is, psychology's scholars apparently have not resolved the basic question of whether or not various complementary or alternative medicines (CAMs) actually *work* from an evidence-based perspective. And so we must return, equivocal and unsatisfied, to our opening question: in the face of such divergent opinions concerning CAM's scientific status, how should professional psychologists regard their relationship to this thriving industry?

We submit that psychology's ambivalence stems at least partly from the fact that CAM's empirical literature contains a vast amount of contradictory data (see Bausell, 2007). Lacking a clear institutional policy statement (e.g., from the APA) on the appropriateness of CAMs for use in psychological practice, inquiring practitioners must face the task of critically appraising this unfamiliar and technically complex literature for themselves; of wading through a seemingly unending sea of academic papers and book chapters touting evidence both for and against the abilities of various CAMs to ameliorate a wide range of health problems. Consider, for instance, the claim that acupuncture—a technique involving the stimulation of specific points on the body, usually via the shallow insertion of sterile needles into the skin—provides reliable pain relief. Despite the publication of at least 3,000 pertinent individual trials (Colquhoun & Novella, 2013), meta-analyses and systematic reviews often reach opposing conclusions (e.g., compare Cherkin et al., 2009; with Madsen, Gøtzsche, & Hróbjartsson, 2009). In this case, even the APA's (2006) *Policy Statement on Evidence-Based Practice in Psychology* breaks down, as it offers no formal recommendations for the resolution of contradictions *within* an evidentiary class, a problem that

[1] See also the Health Service Psychology Education Collaborative's (2013) *Blueprint for Health Service Psychology Education and Training*, which included a call to incorporate CAM into basic training curricula for professional psychologist trainees.

becomes particularly pronounced within the upper tiers of its “best available research” hierarchy (i.e., randomized controlled trials and meta-analyses). Judgment under this sort of uncertainty (too much extra-discipline literature, too many discordant results, and no arbitration protocol) breeds bias (Tversky & Kahneman, 1974), and, we fear, ultimately a field guided more by intuition and ideology than by evidence.

In the present article, we overview two recent intellectual developments that bear directly on this problem of refereeing controversial health intervention research findings. The first, known as the *science-based medicine* movement (Gorski & Novella, 2013), underscores the importance of attending to prior scientific plausibility—the probability that a positive research finding is “genuine” (not the result of statistical artifacts or research design flaws), given the accumulated knowledge of *all* relevant scientific disciplines—when appraising the results of a particular study (e.g., an experiment to determine whether a special kind of massage reduces anxiety). The second development, which we designate the *new statistics and replication* movement in basic psychological science (e.g., Simmons, Nelson, & Simonsohn, 2011) instantiates the first by explicating the myriad ways in which low-quality studies tend to unduly favor false positive results (deciding that a treatment works when it does not), especially when prior scientific plausibility is unestablished.

We will argue that the science-based medicine and new statistics and replication movements’ perspectives can provide professional psychology with a principled and efficient method for sorting strong evidence from weak when making CAM treatment decisions, and, more broadly, render the field’s conflict concerning CAM’s scientific status easier to resolve objectively. To demonstrate, we will apply a simple set of criteria—adapted from scholarship born of both movements; henceforth referred to jointly as the *science-based perspective*—to conduct a systematic re-review of the research cited by Barnett and Shale in their 2012 *Professional Psychology: Research and Practice* contribution as providing empirical support for the effectiveness of 11 different types of CAM that psychologists might consider endorsing. Barnett and Shale’s work has enjoyed a particularly high level of uncritical professional visibility (see, for instance, Barnett & Shale, 2013, featured on the cover of the APA’s *Monitor on Psychology* magazine, and which at the time of this writing may be read by licensed psychologists to earn continuing education credit). Given their potential to influence practice decisions, we believe that an independent test of these authors’ conclusions is necessary.

Perhaps more than any other health service profession, psychology cannot afford simply to ignore CAM’s increasing popularity (Barnes, Bloom, & Nahin, 2008). Estimates suggest that psychotherapy clients utilize CAM at higher rates than the general public (65% versus 40%; Elkins, Marcus, Rajab, & Durgam, 2005), and that they often do so specifically to address mental health problems (Eisenberg et al., 1998). Thus, CAM *directly* impacts psychology’s target clients, and psychologists’ endorsement of CAMs likely contributes to existing public skepticism about psychologists’ role in the United States’ rapidly-evolving health care system (see Lilienfeld, 2012). Moreover, many practicing therapists may already be proffering CAMs to their clients—one survey including more than 600 licensed psychiatrists, psychologists, social workers, and mental health counselors (Brems, Johnson, Warner, & Roberts, 2006) found that 84% reported “suggest[ing] alternative or complementary approaches” *regularly* or *often* in their practice (only 2% had *never* advised CAM use).[2] Whether such suggestions on the part of psychologists warrant commendation (Barnett & Shale, 2012) or correction (Hughes, 2008) depends on psychology’s ability to settle the controversy over CAM’s empirical standing.

---

[2] We suspect that these high endorsement frequencies represent overestimates, perhaps confounded by heterogeneity in practitioners’ conceptions of “alternative or complementary approaches” (a phrase that was undefined and presented without examples in Brems et al.’s study). Nevertheless, these data suggest a non-trivial level of interest.

A Science-Based Perspective

The Science-Based Medicine Movement

History and development. The phrase “science-based medicine” first entered the medical profession’s lexicon formally on January 1st, 2008, when a small group of physicians launched an eponymous Internet web log dedicated to “scientifically examin[ing] medical and health topics of interest to the public” (www.sciencebasedmedicine.org). However, the movement’s conceptual progenitor, *evidence*-based medicine, has served as the *de facto* framework for delivering conventional healthcare since at least the early 1990s (Montori & Guyatt, 2008). Initially, evidence-based medicine represented a proposal for medical pedagogy reform—a model framework for devaluing clinical decision-making strategies predicated on unsystematic observations, intuitive pathophysiological inferences, and deference to authority, each of which had been clearly associated with deleterious patient outcomes. Instead, advocates of evidence-based medicine urged physicians-in-training to consult the research literature when selecting and customizing health interventions for their patients, and, in general, to assign adjudicative priority to the outcomes of scientific experiments.

From a practical standpoint, adopting the evidence-based medicine model meant leveraging the power of the then-nascent Internet, which provided clinicians with on-site access to much larger swaths of published scientific research than any print library could feasibly supply (Montori & Guyatt, 2008). Of course, this newfound access also presented those clinicians with the challenge of picking out the “best” research from an immense and constantly expanding database of empirical reports. Several scholars proffered solutions to this problem in the form of evidence-type hierarchies, often depicted as pyramids (see Figure 1 in this article’s supplemental materials), to help clinicians judge the relative quality and importance of one piece of evidence relative to another (e.g., a case report versus a clinical trial). Typically, systematic summaries of *randomized controlled trials* (RCTs)—experiments in which comparable groups of study participants are assigned randomly to receive different treatments (or no treatment at all)—reside at the top, and, if they are included at all, studies of foundational processes or mechanisms (e.g., *in vitro* models of behavioral mechanisms) fall near the bottom.

In general, these evidence quality pyramids have been successful in aiding healthcare providers who need to compare evidence between major study classes. For instance, when comparing the effectiveness of several different antibiotics, a systematic review is more likely to provide robust evidence than would any single clinical trial, in part because of the systematic review's larger sample size. Similarly, it is unlikely that, when confronted with results from both an RCT and a cohort study, the cohort study (which does not allow for valid inferences about causation) would provide more reliable and valid results. But there are circumstances under which these simple hierarchies fail; circumstances under which even systematic reviews and RCTs will lead science-minded clinicians to spurious conclusions, and to unnecessary patient costs. The *science-based* medicine movement arose explicitly to address such conditions.

A case for prior plausibility. For science-based medicine proponents, the elevation of RCTs on most evidence quality pyramids rests on a critical (but usually unstated) premise: that "lower level" studies, particularly those that focus on suspected pathophysiological mechanisms or basic psychological principles, will have already been performed satisfactorily by the time an RCT report is published (see Gorski & Novella, 2014). That is, in the ideal realization of evidence-based medicine principles, each successful upward-step on the study quality pyramid represents a necessary criterion for advancing to the next, such that the very existence of an RCT for any health intervention should signal that prior plausibility has already been established in basic, proof-of-concept research. In reality, orphan RCTs are quite common, and health researchers are under no particular obligation to demonstrate prior plausibility in a systematic fashion. CAM research seems to violate the upward-progression assumption more so than research on "conventional" treatments (see Novella et al., 2013; Offitt, 2013; and Renckens & Ernst, 2003), typically by offering positive results from a top tier evidentiary class (human clinical trials) *before* conducting studies which might reveal meditating mechanisms.

On the surface, it might seem sensible to begin and end with the "highest quality" research design. That is, if RCTs represent a "gold standard" for obtaining scientific evidence, it ostensibly stands to reason that discovering a statistically and clinically significant difference between randomly assigned treatment and control groups would always trump evidence gleaned from studies lower on the pyramid. Science-based medicine rejects this premise on two grounds.

First, the failure to establish prior plausibility before conducting human clinical trials has led indirectly to wasted research funding, and, more importantly, to patient harm. Take, for instance, a large ($30M) study commissioned to investigate the efficacy of chelation therapy (a technique for removing heavy metals from the body) as a treatment for coronary artery disease, despite the facts that (a) no cogent theories positing a causal role for heavy metal poisoning existed anywhere in the peer-reviewed literature (i.e., no prior plausibility), and that (b) the chelation procedure is risky and invasive (see Nissen, 2013 for a history of these failed trials). Similarly, consider a 2006 Cochrane review that called for more RCTs on the drug Laetrile, a cyanide-based anti-cancer treatment which 20 years earlier was shown to be "toxic and not effective" in basic research (Gorksi, 2014).

Second, even RCTs—endorsed by the APA (2006) as the most effective way to mitigate threats to internal validity—suffer to varying degrees from a host of biases in design and execution. For instance, when control groups (e.g., no-treatment or wait-list conditions) are not equivalent to the treatment group in all respects besides the putative treatment "ingredient," statistical analyses are likely to skew in favor of the research hypothesis, rather than the null. We will return to and expound on this particular claim in a later section of this article. But first, to explicate the nature and magnitude of the alleged false positive problem, we turn now to introduce scientific psychology's burgeoning "new statistics and replication" movement.

Psychology's New Statistics and Replication Movement

History and development. In 2011, Daryl Bem, a prominent social psychologist, claimed to have documented an empirically-verifiable violation of the laws of physics governing our material universe: a *bona fide* case of *precognition*, a purported phenomenon whereby future events can affect people's thinking and behavior in the past through some undefined, paranormal process. For instance, in one of nine experiments Bem described in the esteemed *Journal of Personality and Social Psychology*, research participants had to guess the upcoming position (left or right) of erotic pictures on a computer screen. On average, they guessed correctly 53.1% of the time—a statistically significant result relative to chance (50%)—suggesting to Bem that, somehow, his participants could "feel the future." If vindicated, Bem's results would have overturned centuries of scientific knowledge. Instead, they failed to replicate in independent laboratories (Wagenmakers et al., 2011), leaving research psychologists concerned broadly about the integrity of their empirical literature base. Moreover, according to Wagenmakers and colleagues, there is no indication that Bem engaged in any misconduct—he had played precisely by the same rules that guide all academic publishing in psychological science. A deluge of scholarship on the suspected causes of psychology's "replicability crisis" ensued.

Identifying "researcher degrees of freedom." In a now-landmark subsequent paper, Simmons, Nelson, and Simonsohn (2011) tested an *impossible* hypothesis: that listening to a song about older age ("When I'm Sixty-Four" by The Beatles) can make people (college students) grow younger. They did not predict that students would *feel* younger (perhaps plausible), but rather that they would *actually believe themselves to be chronologically younger* (decidedly implausible). Indeed, the authors found that people on average reported being nearly a year-and-a-half younger after listening to "When I'm Sixty-Four" (M = 20.1 years) than their (randomly assigned and therefore equivalent) control group counterparts (M = 21.5 years) who listened to a non-age-related song, $p < .05$. Thus, Simmons et al. provided a tongue-in-cheek analogue to Bem's precognition studies—a situation in which prior science provides no plausibility for the research hypothesis, but in which standard experimental trials nevertheless found support for it. More importantly, Simmons et al. nominated and described four discrete *causes* of such false positive findings: flexibility in researchers' ability to choose (a) dependent variables, (b) sample size, (c) covariates, and (d) which subsets of experimental conditions to report. Nondisclosure of these behind-the-scenes decisions (deemed *researcher degrees of freedom*), according to the authors, effectively allows social scientists to present anything as statistically significant, and, therefore, as "real." In the view of Simmons et al., researcher degrees of freedom are largely *unconscious*—they capitalize on the instinctive human tendency to seek out information that confirms (rather than information that falsifies) what one *already* believes. Thus, when making routine, seemingly innocuous decisions about study design, data collection, and statistical analyses, generally researchers are unwittingly disposed to err in favor of their hypotheses. And because frequentist statistics—those statistical equations that calculate statistical significance (p) values from a sample to draw conclusions about an entire population—are *inextricably tied to those researcher degrees of freedom*, each additional decision improperly biased in favor of confirming a hypothesis inflates the likelihood of a false positive above and beyond the baseline of 5%—the percentage of studies that, given that there is no genuine effect, *will* produce a false positive result at alpha = .05.

Summary: The Science-Based Perspective

When a study lacks prior plausibility, defined either by a paucity of basic research on an intervention's putative mechanisms of action and/or the invocation of concepts that contradict well-established knowledge from other scientific disciplines, the science-based perspective calls for a commensurately higher standard for acceptance—one which acknowledges the high likelihood that researcher degrees of freedom or some other form of error[3] may have skewed the results of that study toward a false positive conclusion. Thus, when experiments purport to show that some people possess psychic abilities, that listening to a song about old age can cause clocks to run backward, or that inserting needles into a person's skin unblocks the flow of "vital energy," the most likely explanation is a flaw in the research process.

***Ceteris Paribus*: A Science-Based Heuristic for Evaluating CAM Research**

Bausell (2007) was perhaps the first scholar to suggest that these science-based principles could lend themselves to a set of heuristic criteria for evaluating the validity of contentious health intervention studies, and to orphan RCTs in particular. Rather than calling on interested parties to conduct their own scientific literature reviews or to establish a case for or against prior plausibility when considering evidence in favor of a CAM's efficacy, Bausell reasoned that the likelihood of a false positive could be approximated by assessing the quality of a study's *control group*. Specifically, he argued that false positives will occur more often when control groups are not believable from participants' perspectives, and that except for the single active element of the

[3] For instance, academic journals tend to decline the publication of papers that report negative results. This practice seems to be especially pronounced in journals devoted to CAM (Ernst & Pittler, 1997) and in CAM studies conducted in certain countries (Vickers, Goyalb, Harlandb, & Rees, 1998).

treatment believed to cause a specific change, all else in the two interventions should be equal, a concept captured by the Latin phrase *ceteris paribus* (with “other things the same”).

Consider, for instance, a hypothetical study to assess the efficacy of *transcendental meditation*—the practice of repeating personalized mantras to attain special, altered states of consciousness—as a remedy for hypertension. Setting aside the problem of defining and assigning a prior probability level to the claim, assume (a) that a large sample of hypertensive study participants are assigned to either a transcendental meditation or wait-list control group, and (b) that researchers discover a statistically and clinically significant difference ($p < .001$ with a moderate-to-large effect size) between the treatment and control conditions on post-treatment systolic blood pressure readings. Because there are several important systematic differences between the two groups (i.e., little to no *ceteris paribus*), *this study would do nothing to demonstrate that transcendental meditation works* specifically *to reduce hypertension.*

Participants *might* have improved because their unique mantras induced new states of consciousness, but it remains at least equally probable that they improved because their transcendental meditation training engendered many of the nonspecific, anxiety-reducing effects reliably occasioned by the acts of engaging with a compassionate health authority (e.g., a strong client-practitioner relationship), participating in some sort of convincing ritual, and feeling hopeful about improvement (i.e., placebo effects[4]). Coupled with the ever-present threats of researcher degrees of freedom and publication bias, the balance of probability (Occam's razor) strongly favors the latter. Conversely, imagine the same transcendental meditation study modified to include an additional, ritualistically-equivalent control condition—one in which some participants received “personalized” mantras whereas others received a randomly-assigned, pre-fabricated “nonsense” mantra, and in which both groups felt as if they had received a *bona*

[4] See this article’s supplemental materials for a recommended reading list on placebo effects.

fide intervention delivered by a master clinician. Such “dismantling” designs would greatly minimize causes of spurious treatment effects (see Lilienfeld et al., 2014 for a list of 26 such causes) and move closer toward revealing the true, specific effect of the intervention, if there is one.

We submit that requiring higher degrees of *ceteris paribus* in control group designs represents an imperfect but economical way to adjust the bar of evidentiary acceptance in light of prior plausibility. If an intervention really works, if there truly is a “signal in the noise” of researcher degrees of freedom and spurious treatment effects, a preponderance of studies conducted with highly-believable and well-blinded control groups will likely return positive results. Holding CAM research to a standard of *ceteris paribus* at least partially compensates for a lack of prior plausibility at lower, foundational levels on the study-quality pyramid (see Novella et al., 2013) by eliminating all but the specific supposed mechanism-of-action as a cause. Any putative cause that survives deserves serious scientific attention. Thus, *ceteris paribus* represents a simple heuristic that psychologists and laypersons alike can use to make decisions quickly about the credibility of new evidence regarding CAMs.[5]

A Systematic Re-Review of Barnett and Shale (2012)

We turn now to apply the *ceteris paribus* principle—and its pre-requisite criterion that study participants be randomly assigned to either a treatment or control condition—to the selection of empirical research cited by Barnett and Shale (2012) as providing support for the efficacy of several different types of CAM that professional psychologists might consider incorporating into their practice. We believe that Barnett and Shale’s vision for a future in which psychologists endorse CAMs “either through integration into ongoing psychological treatment or through referrals to appropriately trained CAM practitioners” (p. 576) depends critically on their

[5] Cochrane Reviews (www.cochranelibrary.com) often provide detailed and accessible control group descriptions.

ability to defend the position that these CAMs are in fact science-based, particularly given the contentious nature of CAM both within (Hughes, 2008) and outside of (Offitt, 2013) professional psychology. Ethically, the field cannot exempt CAMs from psychology's requirement that professional practice be scientifically well-supported (APA, 2006).

We began our review by retrieving each reference attached to a claim made by Barnett and Shale that a CAM modality has enjoyed empirical support. However, we elected *a priori* to exclude two classes of CAM from our analysis—(a) spirituality (e.g., intercessory prayer), given evidence that such beliefs and behaviors are theoretically and empirically distinct from all other CAMs (Ayers & Kronenfeld, 2010) and that the United States' National Center for Complementary and Integrative Health (NCCIH; 2015) does not include spirituality or prayer in its definitional materials; and (b) dietary supplements, given that Barnett and Shale did not actually recommend that psychologists promote or purvey this variety of CAM. Across the 12 remaining CAM varieties described in the article, we collected a total of 27 references associated with an efficacy claim: three systematic reviews, three meta-analyses, ten RCTs, four within-subject designs, two non-randomized efficacy comparisons, two case-studies, one qualitative systematic review, one trade book, and one link to a promotional website.

RCTs. First, we turned our attention to Barnett and Shale's individual RCT citations (10 in total). Seven failed to achieve an acceptable degree of *ceteris paribus* (see Table 1 for control group descriptions). Two of the remaining three did not actually report statistically significant differences between treatments and *ceteris paribus* control groups.

Systematic reviews and meta-analyses. Next, we assessed the degree of *ceteris paribus* in each of the individual studies (N = 229) aggregated by the six systematic reviews and meta-analyses. See Table 2 (supplementary materials) for (a) control group descriptions and (b) explications of our *ceteris paribus* criteria within each CAM category (i.e., what we consider to be a reasonable *ceteris paribus* design for different types of CAM). Less than 15% employed research designs that we deemed soundly capable of ruling out non-specific effects, and more than half of these failed to detect statistically significant differences relative to the *ceteris paribus* control group (see Table 2, supplemental, for full details).

Conclusions. From a science-based perspective, the highest-quality empirical studies (RCTs) cited by Barnett and Shale as evidence for CAM's efficacy in fact support the opposite conclusion—that CAMs either have not been properly evaluated or simply do not work better than credible placebos. Of course, the collection of studies that we considered here clearly represents a small, biased sampling of the relevant research literature. However, that bias should have *positively* skewed the results, given that (a) Barnett and Shale apparently did not articulate and therefore apparently did not adhere to any systematic protocol for choosing their supporting references, and (b) the aim of their article was ostensibly to promote, rather than to critically evaluate the use of CAMs. Perhaps future research will discover exceptions, but, given the agreement of our findings with the conclusions of other scholars who have approached the issue of CAM efficacy though a similar (science-based) lens, the burden of proof would seem to lie now with those who make extraordinary claims in support of implausible CAMs.

One might sensibly question the logic of choosing "CAM," writ large, as our unit of analysis—the category consists of a vast, diverse, and ever-burgeoning catalog of products and practices, some of which surely are more scientifically plausible than others. Biofeedback, for instance, operates under the assumption that autonomic regulatory processes (such as heart rate or core body temperature) can be controlled via top-down (conscious) cognitive effort, and that providing patients with "feedback" about those autonomic processes (e.g., by showing patients a real-time graph depicting heart-rate data) allows them to take mental control (e.g., by concentrating their thoughts to reduce their heart rate). *Prima facie*, no basic scientific principles need be re-written for the biofeedback-control hypothesis to be accepted as a genuine phenomenon. However, where there is a specific claim, there should be a specific test, and to date relatively few biofeedback RCTs have utilized the strongest *ceteris paribus* design we can imagine: providing control group participants with bogus but believable feedback. The same is true for the panoply of relaxation techniques that are classified as CAM by the NCCIH (2015). There may be viable hypotheses, but until a proposed treatment has established prior plausibility in basic research and passed the *ceteris paribus* test, it must remain in question and therefore outside of psychology's scope of evidence-based practice.[6]

What's the Harm?

One might argue that the science-based perspective imposes unnecessary restrictions on evidence-based practice in psychology. Indeed, the preponderance of evidence from studies that we reviewed shows clearly that using CAM will result in better patient outcomes than would providing no treatment at all, and advocates often claim that CAM carries few if any injurious side effects. Is there any harm in offering an intervention that produces only nonspecific salutary effects? We will argue on four grounds that the answer to this question is *yes*, there is harm.

First, CAMs are not risk-free. Dietary supplements and "herbal" products are currently unregulated in the United States, and many—estimates range up to 33% of all herbal products available to consumers (Newmaster, et al., 2013)—have been found to contain active and potentially dangerous ingredients that do not appear on their product labels. Acupuncture needles have caused painful bacterial infections (Song et al., 2006), chiropractic manipulations have left

[6] Given that many psychotherapy clients are already using CAMs, professional psychologists might consider advocating for more the large-scale funding of rigorous, double-blind, controlled studies with large, representative samples.

patients with permanent spinal injuries (Ernst, 2007), and megavitamin therapies have poisoned many unsuspecting consumers (Timbo et al., 2006). Even if the overall base rates for such adverse events are low, these CAMs' risks still outweigh their potential benefits when other, less hazardous, science-based methods for achieving palliation through nonspecific effects (e.g., common-factors-based psychotherapy) are available.

Second, CAMs harm clients by imposing opportunity costs. That is, by utilizing CAMs, clients may delay or forego the opportunity to utilize more efficacious (and empirically-validated) treatments (e.g., see Benmeir et al., 1991; Carllson, Bergqvist, & Hellgren, 1996; Hainer et al., 2000; Luyckx, Steenkamp, & Stewart, 2005; Malik & Gopalan, 2003; and Vohra, Johnston, Cramer, & Humphreys, 2006). A Reiki practitioner is unlikely to cause direct physical harm to a client through the ceremonial waving of hands (from a safe distance, of course), unless that client chose to place his or her faith in the Reiki master *instead* of a conventional healthcare provider. Such opportunity costs can be fatal when the client's presenting ailment is life-threatening.

Third, psychologists who purvey CAMs may be violating clients' rights to autonomous decision-making. We can envision two varieties of clinicians who might endorse CAMs: (a) those who are unfamiliar with or who have misinterpreted the scientific evidence—we have argued that such unfamiliarity or misinterpretation represents the essential trouble with CAM in professional psychology—and (b) those who would deem it acceptable to endorse a treatment that evokes only placebic responses (i.e., placebo medicine). In our view neither position is ethically or professionally defensible. Practitioners who recommend CAMs are by definition deviating from the generally-accepted standard of care,[7] which does not include CAMs, and

[7] When they adopt the Scientist-Practitioner model, graduate training programs commit to inculcating in all their professional psychology trainees a commitment to use this evidence-based standard (Baker et al., 2009).

therefore they must personally shoulder the burden of demonstrating efficacy and safety. The results of our systematic re-review suggest that, from a science-based perspective, they cannot reasonably make this case. Those who would prescribe CAMs knowing that the evidence does not support their efficacy may have failed by omission to practice with integrity (APA Ethics Code, 2010; Aspirational Principle C).

When clients' CAM preferences arise from cultural influences, psychologists can both remain multiculturally-sensitive *and* uphold standards of evidence-based practice by discussing the fact that psychologists espouse a *scientific* epistemology, but that other practitioners (such as indigenous folk healers) may select their interventions using other ways of knowing. Skepticism need not legitimize disrespect or intolerance if psychologists engage earnestly with their clients regarding their belief systems. In this way, belief in CAM resembles religion/spirituality, and it can be approached in a similar fashion (see Plante, 2007).

Fourth, many laypeople already perceive the field of psychology as unscientific (Lilienfeld, 2012). Consider, for example, *Newsweek* science editor Sharon Begley's 2009 article, "Ignoring the Evidence: Why Do Psychologists Reject Science?" Citing recent calls in the psychological literature for increased scientific rigor in the training of clinical psychologists, Begley lamented what she perceived as a rocky relationship between current mental health practices and scientific research. Allowing professional psychology to embrace CAM, and, as a consequence, to shift away from the scientific mainstream, will only exacerbate this problem. It would blur the lines between science and pseudoscience in the profession, and almost certainly thwart the many impressive efforts psychologists have undertaken to secure full parity for reimbursement of non-pharmacological mental health services in managed care. Neither mental health professionals, nor the public they are committed to serving, can afford such a setback.

A Science-Based Vision for the Future of Professional Psychology

We believe that the science-based perspective represents a promising future for professional psychology. Its broad adoption will accelerate psychologists' search for the most effective methods of helping those in distress by helping practitioners to sort the wheat from the chaff in healthcare; to distinguish between treatments that work, treatments that appear to work but wither under the lights of controlled observation (CAMs), and treatments that do measurable harm (also CAMs). From the science-based perspective, the preponderance of evidence points to an unambiguous answer to the question that we posed at the beginning of this article: given the current state of the evidence, practicing psychologists should reject CAM.

References

American Psychological Association Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist, 61*, 271–285. http://dx.doi.org/10.1037/0003-066X.61.4.271

American Psychological Association (2010). *Ethical principles of psychologists and code of conduct*. Retrieved from http://www.apa.org/ethics/code2002.html

Ayers, S. L., & Kronenfeld, J. J. (2010). Using factor analysis to create complementary and alternative medicine domains: An examination of patterns of use. *Health*, *14*(3), 229–52. doi:10.1177/1363459309347491

Baker, T. B., McFall, R. M., & Shoham, V. (2009). 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. PMID: 20865146

Barnes, P. M., Bloom, B., & Nahin, R. L. (2008). Complementary and alternative medicine use among adults and children: United States, 2007. *National Health Statistics Reports, 12,* 1–23. PMID: 19361005

Barnett, J. E., & Shale, A. J. (2012). The integration of Complementary and Alternative Medicine (CAM) into the practice of psychology: A vision for the future. *Professional Psychology: Research and Practice, 43(6),* 576–585. doi: 10.1037/a0028919

Barnett, J. E., & Shale, A. J. (2013). Alternative techniques. *Monitor on Psychology, 44(4),* 48–56. Retrieved October 4, 2014 from http://www.apa.org/monitor/2013/04/ce-corner.aspx

Barnett, J. E., Shale, A. J., Elkins, G., & Fisher, W. (2014). *Complementary and alternative medicine for psychologists: An essential resource*. Washington, DC: American Psychological Association.

Bassman, L. E., & Uellendahl, G. (2003). Complementary/alternative medicine: Ethical, professional, and practical challenges for psychologists. *Professional Psychology: Research and Practice, 34(3)*, 264–270. doi:10.1037/0735-7028.34.3.264

Bausell, R. B. (2007). *Snake oil science: The truth about complementary and alternative medicine*. New York, NY: Oxford University Press.

Begley, S. (2009, October). Ignoring the evidence: Why do psychologists reject science? *Newsweek.* Retrieved from http://www.newsweek.com/id/216506

Bem, D. J. (2011). Feeling the future: experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology, 100(3)*, 407–425. doi:10.1037/a0021524

Benmeir, P., Neuman, A., Weinberg, A., Sucher, E., Weshler, Z., Lusthaus, S., Rotem, M,m Eldad, A., & Wexler, M. R. (1991). Giant melanoma of the inner thigh: A homeopathic life-threatening negligence. *Annals of Plastic Surgery, 27(6)*, 583-585. PMID: 1793246

Brems, C., Johnson, M. E., Warner, T. D., & Roberts, L. W. (2006). Patient requests and provider suggestions for alternative treatments as reported by rural and urban care providers. *Complementary Therapies in Medicine, 14(1)*, 10–19. PMID: 16473749

Carllson, T., Bergqvist, L., & Hellgren, U. (2006). Homeopathic resistant malaria. *Journal of Travel Medicine, 3(1)*, 62. doi: 10.1111/j.1708-8305.1996.tb00700.x

Cherkin, D. C., Sherman, K. J., Avins, A. L., Erro, J. H., Ichikawa, L, Barlow, W. E., … Deyo RA. (2009). A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain. *Archives of Internal Medicine, 169(9)*, 858–866. doi: 10.1001/archinternmed.2009.65.

Colquhoun, D., & Novella, S. P. (2013). Acupuncture is theatrical placebo. *Anesthesia &*

Analgesia, 116(6), 1360–1363. doi: 10.1213/ANE.0b013e31828f2d5e

Dale, R. A. (1994). The forms of Qi (vital energy) in acupuncture. American Journal of Acupuncture, 22(1), 71-76. Retrieved from http://www.acupuncture.com/education/theory/formsqi.htm

Eisenberg, D. M., Davis, R. B., Ettner, S. L., Appel, S., Wilkey, S., Van Rompay, M., & Kessler, R. C. (1998). Trends in alternative medicine use in the United States, 1990-1997: Results of a follow-up national survey. *JAMA*, *280(18)*, 1569–1575. PMID: 9820257

Elkins, G., Marcus, J., Rajab, M. H., & Durgam, S. (2005). Complementary and alternative therapy use by psychotherapy clients. *Psychotherapy*, *42(2)*, 232–235. doi:10.1037/0033-3204.42.2.232

Ernst, E. (2007). Adverse effects of spinal manipulation: A systematic review. *Journal of the Royal Society of Medicine*, 100(7), 330-338. http://dx.doi.org/10.1258/jrsm.100.7.330

Ernst, E., & Pittler, M. H. (1997). Alternative therapy bias. *Nature*, 385(6616), 480. 9020351

Gorski, D. H. (2014). Integrative oncology: Really the best of both worlds? *Nature Reviews Cancer*. Published online 18 September 2014. doi:10.1038/nrc3822

Gorski, D. H., & Novella, S. P. (2014). Clinical trials of integrative medicine: Testing whether magic works? *Trends in Molecular Medicine, 20(9),* 473–476. PMID: 25150944

Hainer, M. I., Tsai, N., Komura, S. T., & Chiu, C. L. (2000). Fatal hepatorenal failure associated with hydrazine sulfate. *Annals of Internal Medicine, 133(11),* 877-880. doi:10.7326/0003-4819-133-11-200012050-00011

Health Service Psychology Education Collaborative. (2013). Professional psychology in health care services: A blueprint for education and training. *American Psychologist, 68*, 411–426. doi:10.1037/a0033265.

Hughes, B. M. (2008). How should clinical psychologists approach complementary and alternative medicine? Empirical, epistemological, and ethical considerations. *Clinical Psychology Review, 28(4),* 657–75. doi:10.1016/j.cpr.2007.09.005

Hymel, G. M., & Rich, G. J. (2014). Health psychology as a context for massage therapy: A conceptual model with CAM as mediator. *Journal of Bodywork and Movement Therapies, 18(2)*, 174–182. doi:10.1016/j.jbmt.2013.07.001

Lilienfeld, S. O. (2012). Public skepticism of psychology: Why many people perceive the study of human behavior as unscientific. *American Psychologist*, *67*(2), 111–29. PMID: 21668088

Lilienfeld, S. O., Ritschel, L. A., Lynn, S. J., Cautin, R. L., & Latzman, R. D. (2014). Why ineffective psychotherapies appear to work: A taxonomy of causes of spurious therapeutic effectiveness. *Perspectives on Psychological Science*, *9(4)*, 355–387. doi:10.1177/1745691614535216

Luyckx, V. A., Steenkamp, V., & Stewart, M. J. (2005). Acute renal failure associated with the use of traditional folk remedies in South Africa. *Renal Failure, 27(1)*, 35-43. PMID: 15717633

Madsen, M., Gøtzsche, P., & Hróbjartsson, A. (2009). Acupuncture treatment for pain: Systematic review of randomised clinical trials with acupuncture, placebo acupuncture, and no acupuncture groups. *BMJ*, *338,* 1-8. doi: http://dx.doi.org/10.1136/bmj.a3115

Malik, I. A., & Gopalan, S. (2003). Use of CAM results in delay in seeking medical advicee for breast cancer. *European Journal of Epidemiology, 18(8)*, 817-822. PMID: 12974558

Montori, V. M., & Guyatt, G. H. (2008). Progress in evidence-based medicine. *JAMA, 300(15),* 1814–1816. doi:10.1001/jama.300.15.1814

National Center for Complementary and Integrative Health. (2015). *Complementary, Alternative,* *or Integrative Health: What's In a Name?* https://nccih.nih.gov/health/integrative-health

Newmaster, S. G., Grguric, M., Shanmughanandhan, D., Ramalingam, S., & Ragupathy, S. (2013). DNA barcoding detects contamination and substitution in North American herbal products. *BMC Medicine, 11(1),* 222. doi:10.1186/1741-7015-11-222

Nissen, S. E. (2013). Concerns about reliability in the Trial to Assess Chelation Therapy (TACT). *JAMA, 309(12),* 1293–1294. doi:10.1001/jama.2013.2778

Novella, S. P., Gorski, D. H., & Crislip, M. A. (Eds.) (2013). *Science-based medicine: "Best Of" anthology*. Falls Church, VA: James Randi Educational Foundation.

Offit, P. A. (2013). *Do you believe in magic? The sense and nonsense of alternative medicine*. New York, NY: Harper.

Park, C. L. (2013). Mind-body CAM interventions: Current status and considerations for integration into behavioral medicine. *Journal of Clinical Psychology, 69,* 45 -63. http://dx.doi.org/10.1002/jclp.21910

Plante, T. G. (2007). Integrating spirituality and psychotherapy: Ethical issues and principles to consider. *Journal of Clinical Psychology, 63(9),* 891-902. doi: 10.1002/jclp.20383

Renckens, C. N. M., & Ernst, E. (2003). Some complementary and alternative therapies are too implausible to be investigated. *Focus on Alternative and Complementary Therapies, 8(3),* 307-308. doi: 10.1211/fact.2003.00304

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22(11),* 1359–1366. http://dx.doi.org/10.1177/0956797611417632

Song, J. Y., Sohn, J. W., Jeong, H. W., Cheong, H. J., Kim, W. J., & Kim, M. J. (2006). An outbreak of post-acupuncture cutaneous infection due to Mycobacterium abscessus. *BMC*

Infectious Diseases, 6, 6. doi:10.1186/1471-2334-6-6

Timbo, B. B., Ross, M. P., McCarthy, P. V, & Lin, C. T. J. (2006). Dietary supplements in a national survey: Prevalence of use and reports of adverse events. *Journal of the American Dietetic Association, 106(12),* 1966–1974. doi:10.1016/j.jada.2006.09.002

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185(4157),* 1124–1131. doi:10.1126/science.185.4157.1124

Vickers, A., Goyal, N., Harland, R., & Rees, R. (1998). Do certain countries produce only positive results? A systematic review of controlled trials. *Controlled Clinical Trials, 19(2),* 159–166. http://dx.doi.org/10.1016/S0197-2456(97)00150-5

Vohra, S., Johnston, B., Cramer, K., & Humphreys, K. (2006). Adverse events associated with pediatric spinal manipulation: A systematic review. *Pediatrics, 119(1),* e275-e283. doi: 10.1542/peds.2006-1392

Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology, 100(3),* 426–432. doi:10.1037/a0022790

White, K. P. (2000). Psychology and complementary and alternative medicine. *Professional Psychology: Research and Practice, 31(6),* 671–681. doi:10.1037//0735-7028.31.6.671