Professional Psychology: Research and Practice
2004, Vol. 35, No. 6, 563–570

Copyright 2004 by the American Psychological Association
0735-7028/04/$12.00 DOI: 10.1037/0735-7028.35.6.563

Attending to the Omissions: A Historical Examination of Evidence-Based Practice Movements

Bruce E. Wampold and Kuldhir S. Bhati
University of Wisconsin—Madison

Evidence-based practice and empirically supported treatment movements are potent forces that affect the practice of psychology today and have the potential to mandate the types of treatments psychologists conduct. The histories of these movements reveal that certain aspects of therapy valued by psychologists have been ignored. It is shown that the evidence-based movements (a) overemphasize treatments and treatment differences and (b) ignore aspects of psychotherapy that have been shown to be related to outcome, such as variation among psychologists, the relationship, and other common factors. It is important that psychologists understand the development of these movements so that they can be critical consumers of research and can effectively influence the future course of events.

The practice of psychology is increasingly being affected by the systems that pay for psychological services. As a field, psychology has responded to these pressures by demonstrating that various psychological treatments are based on empirical evidence, which in turn has led to attempts to mandate the types of treatment provided by psychologists. However, many practitioners have resisted the efforts by managed care and some academic psychologists to alter their practice. In contentious contexts, it is often helpful to understand where things originated and to be cognizant of the forces that furthered various aspects while limiting others. The purpose of this article is to provide a historical perspective of the development of the evidence-based practice movement, highlighting the omissions that have occurred so as to enable psychologists to advocate more effectively in the debates that are shaping practice. This perspective will strengthen our role as scientist-practitioners in the broadest sense.

History and progress are constructed and defined by events and choices made. Foucault (1965) termed the social sciences, including psychology and medicine, as the *human sciences* to emphasize and highlight the fact that they are constructed through the actions and choices of various actors. Looking backward through the lens of progressivism, a seemingly inevitable march toward enlightenment is revealed. The occasional blemish (e.g., in the social sciences, the eugenics movement) is either ignored or characterized as an anomaly. This retrospective approach typically leads to an enthusiastic embrace of contemporary events as steps forward; critics are characterized as nostalgic and anachronistic at best or pessimistic and obstructionist at worst. However, a reasoned examination of history, noting the omissions as well as the events, is critical if one is to have a balanced contextual understanding of current trends.

When one looks at contemporary scientific developments in the mental health field, two movements predominate: evidence-based practice (EBP), which originated in the United Kingdom, and empirically supported treatments (ESTs), which originated in the United States (where appropriate, we consolidate them together under the rubric of *evidence-based movements*). In this article, we present both a retrospective and a prospective examination of the evidence-based treatments, elucidating the changes that have occurred and revealing how progress shaped the omissions. In this way, the costs as well as the benefits of EBP can be properly assessed from the perspective of psychological practice.

The Retrospective

Evidence-Based Practice

In medicine, the EBP movement, begun recently in the United Kingdom, reflects the desire to base medical practice on the laudable goal of optimizing patient outcomes by translating the evidence derived from research into practice. This movement reflected the confluence of the development of science and modern medicine, the use of particular research methods, and the pressure to be accountable for outcomes, which is exerted by mental health financial systems. The advocates of EBP cite numerous examples of the harm created by failing to heed accumulating evidence. One perspicuous example involves the use of streptokinase, an enzyme that dissolves clots, as a treatment for acute myocardial infarction (AMI; Hunt, 1997). Clinical trials began as early as 1959, but the results, owing to small sample sizes, were inconclusive—some supported the use of streptokinase, others not. However, had the

BRUCE E. WAMPOLD received his PhD from the University of California, Santa Barbara, and is currently professor and chair of the Department of Counseling Psychology at the University of Wisconsin—Madison. He is a fellow of the American Psychological Association and a diplomate of the American Board of Professional Psychology (counseling psychology). His research interests are in investigating the empirical, historical, and anthropological aspects of psychotherapy as well as quantitative methods in applied psychology.

KULDHIR S. BHATI received his MA in counseling psychology from Boston College and is currently a student in the PhD program in counseling psychology at the University of Wisconsin—Madison. His areas of interest are the history and philosophy of science, particularly as they relate to the development of psychotherapy as a profession.

CORRESPONDENCE CONCERNING THIS ARTICLE should be addressed to Bruce E. Wampold, Department of Counseling Psychology, University of Wisconsin—Madison, 321 Education Building, 1000 Bascom Mall, Madison, WI 53706-1398. E-mail: wampold@education.wisc.edu

results been accumulated by meta-analysis as early as 1969,[1] the efficacy of this treatment would have been established. Iain Chalmers, an early advocate of meta-analyses, made the following observation:

> Streptokinase was the classic example. The meta-analyses showed clearly that the effect on mortality was statistically significant, but the experts in cardiology and the textbook authors whose opinions dominated the field weren't even beginning to recommend it until the late 1980s, and then only little by little. (see Hunt, 1997, p. 87)

Finally, in 1986 and 1988, large-scale clinical trials of the type acceptable to the U.S. Food and Drug Administration (FDA) were conducted, and subsequently the treatment was generally accepted. However, in the meantime, it is estimated that tens of thousands of patients died because they were not administered streptokinase. It also demonstrates how meta-analysis, a method critical to the EBP rationale, synthesizes evidence in a way that can inform medicine and save lives and that is superior to simply looking at the results of individual clinical trials. Although research evidence has been a hallmark of modern medicine, the EBP movement emphasizes systematic and analytic reviews of evidence and the examination of practice in light of the results of such reviews.

The EBP of medicine has been extended to the practices in the mental health field in the United Kingdom, heralded by the launching of a new journal, *Evidence-Based Mental Health* (Geddes, Reynolds, Streiner, & Szatmari, 1997), in 1998. The aim of this journal, published by the BMJ (formerly *The British Medical Journal*) in collaboration with the Royal College of Psychiatrists, the British Psychological Society, and the Royal College of Nursing, is to "provide all mental health clinicians with the very best information about mental health care in the form of 'value added' abstracts" (Geddes et al., 1997, p. 1484).

Empirically Supported Treatments

In the United States, the emphasis on using research to guide clinical practice is best exemplified by the EST movement, which began in 1993 with Division 12 of the American Psychological Association's (APA) Task Force on Promotion and Dissemination of Psychological Procedures (1995). The purpose of this task force, and subsequent updates and refinements (Chambless et al., 1998), was to identify and disseminate treatments that have been demonstrated, through rigorous research, to be efficacious treatments for specific disorders. Although the criteria for identifying such treatments have been modified over the years to address various concerns and to account for developments in research design, the essence of the criteria (see the Appendix for a recent version; Chambless et al., 1998) are based on the FDA criteria for approving drugs.

FDA and Clinical Trials

Essentially, FDA rules for approving drugs are based on the randomized, double-blinded placebo control group design. Drugs are approved if they can be shown in a number of trials to be superior to a placebo, which is designed to be indistinguishable from the active medication; the logic is that the specific ingredients of the active medication are responsible for benefits over and above psychological effects such as hope, expectation, remoralization, and so forth. The origins of this design are found in the 1950s, but it was not until 1980 that such designs were required by the FDA for drug approval (Shapiro & Shapiro, 1997).

Although the randomized placebo control group design is so well established that we take its logic for granted, the history of the development of the design is one of remarkable accomplishment achieved through contribution from diverse sources. Danziger (1990; see also Wampold, 2001a) discussed how the confluence of three strands resulted in the origins of the randomized control group design. The first strand was Wilhelm Wundt's development of the experimental method in psychology. In Wundt's laboratory, in the second half of the 19th century, he and his students, in the search for general laws, exchanged the roles of experimenters and observers; the experimenters would arrange various situations and the observers would report their perceptions as experts, in much the way that physicists of the time observed experiments and reported results. The second strand was related to the measurement of mental abilities in Britain by such luminaries as Sir Francis Galton and Karl Pearson, in the late 19th century, and involved the development of measurements and statistics that characterized distributions of traits so that normality (i.e., the central tendency of the distribution) and deviations from normality (i.e., the tails of the distributions) could be defined and people classified accordingly. The third strand emanated from France, at about the same time as Wundt was conducting his laboratory experiments, and involved physicians testing the effects of various treatments, particularly hypnosis, through comparisons between groups of patients. Wundt provided the experimental method, the French the comparison of groups in which participants were subjected to treatments, and the British the statistical means to characterize the responses of participants.

What were needed next were the concepts of randomization, random error, and comparison of systematic variance with random error (i.e., the analysis of variance). These ideas were contributed by many, but most notably by Sir Ronald A. Fisher (see, e.g., Fisher, 1935). Fisher's publication of *The Design of Experiments* (1935) was pivotal as it became the design and statistical bases of clinical trials that were used by medical researchers to show the efficacy of various medications. The final step was the development of the placebo control. This critical addition to the randomized design, discussed by Fisher and others, was developed in the United States and the United Kingdom, beginning in the 1930s and continuing through the 1950s and 1960s (Gehan & Lemak, 1994; Shapiro & Shapiro, 1997; Wampold, 2001a), and it introduced the concept of placebo control for the purpose of holding constant all factors except the medication's active ingredient to establish that the benefits of medications were due to physiochemical properties rather than to patients' expectations, hopes, or other psychological processes. The validity of this design requires that the placebo pill must be indistinguishable from the active medication—in appearance, taste, and smell—and the patient, the administrator of the

[1] Although methods for quantitatively aggregating various statistics across studies existed in the literature for decades prior to the 1970s, it was Gene Glass who popularized meta-analysis in the 1970s by his explanation and application of the method to examine psychotherapy efficacy and the effect of classroom size on achievement (Hunt, 1997).

treatment, and the evaluator must be unaware of the patient's treatment condition (i.e., the experiment is double blinded).

Placebos and Anton Mesmer

Why was the placebo control so important to modern medicine and sine qua non for FDA approval of drugs? One answer is found embedded in the story of Anton Mesmer. Mesmer was by all accounts an extremely charismatic man who by training was a physician. In his dissertation (Mesmer, 1766/1980), Mesmer claimed that some illnesses arose from the disruption of the normal flow of an invisible universal fluid, which he called *animal magnetism*. A well-trained physician could learn to locate the blocks (those causing the disruption of the flow of the fluid) and by touch, massage, and so on remove the blocks and rechannel animal magnetism through the body, thereby curing the patient (Buranelli, 1975; Gauld, 1992). After further "research," Mesmer found that he could "magnetize" objects with animal magnetism and these could be used to cure his clients, which rather fortuitously enabled him to offer his services to many more people (Gallo & Finger, 2000; Pattie, 1994). The success of this treatment was well documented and led to its immense popularity in the late 18th century.

The late 18th century was a time of great transitions; science was struggling with theology for legitimacy, wondrous new things were being discovered (e.g., gravity, electricity), and all sorts of "treatments" were being peddled to people. Shapiro and Shapiro (1997) tallied 4,785 drugs and 16,842 prescriptions used in antiquity; the number of remedies grew and by Mesmer's time included such wonderful substances that mixed ingredients such as moss from the skull of victims of violent death, viper's flesh, live frogs, worms, crab eyes, powders of precious stones, scorpions, spider webs, wood lice, saliva of fasting humans, and sexual organs. With precious few exceptions, these substances had no specific benefits, a fact that was becoming embarrassingly obvious at the dawn of the scientific revolution. The term *placebo* was used to denote substances given to please or placate patients.

It was time to expose a charlatan who used methods deemed unscientific despite the benefits that might accrue to patients. And there was no better target that the immensely popular Mesmer. Consequently, in 1784, King Louis XVI of France established a Royal Commission, chaired by Benjamin Franklin and consisting of such luminaries as Antoine Lavoisier (the father of modern chemistry) to investigate mesmerism (Gould, 1991). The noted natural historian Stephen Jay Gould has heralded the testing and discrediting of Mesmer as one of the earliest and an exemplary instance in which the scientific method was used to expose pseudo-science and charlatanism (Gould, 1989). The Royal Commission designed a number of experiments to test the existence of animal magnetism and irrefutably showed that Mesmer's theory was not valid, in that the existence of the fluid and its flow was not proved (Pattie, 1994). Some of the experiments involved the patients being split into two groups, with one group coming into contact with "magnetized" objects and the other group coming into contact with placebos (what they believed were "magnetized" objects), so as to test if their reactions ("crises") were due to the treatment or due to the powers of suggestion. This enabled the Royal Commission to demonstrate that the cure did not occur through the treatment-specific ingredients and it emphasized the benefits of utilizing placebo controls to test treatments.

And so we have seen the transition from healing by magnetized objects to administering streptokinase for AMIs; from the use of wild concoctions of every exotic type of ingredients to the use of only those substances established by rigorous clinical trials; from the practice of healing using any method deemed reasonable to the practitioner to the use of procedures that are evidence based or empirically supported. Progress indeed. However, an examination of the history prospectively will show some of the costs incurred in the building of the evidence-based edifice.

The Prospective

Mesmer Revisited

Mesmer was certainly not a charlatan in that there was no intent to deceive or defraud. His theories of bodily action were extensions of the emerging theories of physics current at the time. Sir Isaac Newton postulated gravity as a force between objects and showed how the gravity of the moon and sun formed the tides; Mesmer postulated how gravity affected the fluids in the body and similarly adapted theories of magnetism. Mesmer dabbled in the occult, but then for decades Newton was similarly immersed in the supernatural and alchemy—numerous notebooks filled with copious notes on such things are scattered among his early scientific work (Gleick, 2003). To comprehend Mesmer and his treatments, we must keep in mind that he lived in a period dominated by theories of the supernatural and alchemy; Sir Isaac Newton was able to break free of such musings, but he was unique. Mesmer was a serious scientist of the times, a member of the Bavarian Academy of Sciences, and a sought-after physician by the French aristocrats (Mozart regularly performed at his residence). The commissioners acknowledged the effects of Mesmer's treatment: "It is impossible not to recognize in these regular effects an extraordinary influence acting upon the patients, making itself master of them, and of which he who super-intends the process appears to be the depositary" (Walmsley, 1967, p. 134; see also Franklin, 1785). Nevertheless, Mesmer's treatment failed the emerging "specificity" test: Does the treatment work through its hypothesized mechanism?

Specificity in Medicine and Psychotherapy

The case of Anton Mesmer demonstrates the need for specificity—his treatment was effective, but not through the specific ingredients he hypothesized. Anton Mesmer was discredited because the Royal Commission could not verify the existence of animal magnetism and therefore invalidated the basis of the treatment, despite its recognition that the treatment was remarkably beneficial. Modern medicine is, as we have seen, dependent on demonstrating specificity; indeed, specificity is the central premise of Western medicine. Without specificity, medicine would have to retreat to the large pharmacopeias of Mesmer's time, as many of these remedies were effective by the power of the placebo. As we have seen, medicine invented the randomized double-blind control group design to demonstrate that the benefits of substances and procedures were due to the specific ingredients and not to other factors, particularly psychological factors. In medical studies, rarely is a no-treatment control used, as the effectiveness of a treatment vis-à-vis no treatment is not convincing evidence that the

treatment deserves recognition within the philosophy of medicine. Rather a treatment must show a distinct (i.e., statistically significant) advantage over what could be achieved through the various psychological processes.

Shortly after the randomized double-blind placebo control group design was developed in medicine, Rosenthal and Frank (1956) suggested that the design be used in psychotherapy research to rule out factors that are common to therapies and to establish specificity:

> It may be possible to study the possible specific effects of any particular form of therapy by the use of a matched control group participating in an activity regarded therapeutically inert from the stand point of the theory of the therapy being studied. That is, it would not be expected to produce the effects predicted by the theory. The "placebo psychotherapy" in a sense would be analogous to placebos in that it would be administered under circumstances and by persons such that the patients would be expected to be helped by it. (pp. 299–300)

According to this logic, if cognitive–behavior therapy (CBT) for depression were compared with an adequate placebo control group and found to produce superior outcomes, then it could be concluded that the purported active ingredients in CBT (e.g., altering core schema and challenging irrational thoughts) were responsible for the benefits of the treatment. Unfortunately, there are two essential problems with placebo control groups in psychotherapy that attenuate their validity to establish that the benefits of psychotherapy are derived from the specific ingredients (Basham, 1986; Baskin, Tierney, Minami, & Wampold, 2003; Borkovec & Nau, 1972; Brody, 1980; P. Horvath, 1988; O'Leary & Borkovec, 1978; Shepherd, 1993; Wampold, 1997, 2001a, 2001b). First, psychotherapy studies cannot be blinded in the manner used in placebo-controlled medical studies. Quite obviously, the psychologist must be aware of the treatment being delivered to follow the treatment protocol (O'Leary & Borkovec, 1978; Seligman, 1995; Wampold, 2001b). Consequently, the psychologist, delivering a treatment that is designed to be a sham (a technical term often used to denote the placebo condition), will likely believe that the placebo treatment will be ineffective and will produce cues consistent with the attenuated faith in the treatment, whether blatant or subtle, which unavoidably will be communicated to the patient. Second, the placebo is not indistinguishable from the active treatment, and the apparent differences between placebo and active treatment play havoc with the logic of the design. For example, consider the following comparison between an active treatment (interpersonal therapy) and a placebo (supportive psychotherapy), which purportedly was used to establish the specificity of interpersonal therapy:

> Supportive psychotherapy, defined as noninterpersonal psychotherapy and noncognitive behavioral therapy, resembles the client-centered therapy of Rogers, with added psychoeducation about depression and HIV. Unlike interpersonal psychotherapists, supportive psychotherapists offered patients no explicit explanatory mechanism for treatment effect and did not focus treatment on specific themes. Although supportive psychotherapy may have been hampered by the proscription of interpersonal and cognitive techniques, it was by no means a nontreatment, particularly as delivered by empathic, skillful, experienced, and dedicated psychologists. Sixteen 50 minute sessions of interpersonal therapy were scheduled within a 17-week period. The supportive psychotherapy condition had between eight and 16 sessions, determined by patient need, of 30–50 minute duration. (Markowitz et al., 1995, p. 1505)

It is clear that, in this design, the control is inadequate because the dose of therapy (i.e., the amount, in number of sessions and length of session) was smaller in the control than in the treatment, no credible rationale was provided to patients in the control group, and the psychologists in the control condition were aware that they were delivering a sham treatment.

Psychotherapy placebos do not have the same experimental validity that they have in medical context; their flaws put them at a disadvantage to the active treatment. Nevertheless, placebo-type controls (sometimes called common factor controls, alternative treatments, and so forth) are used by researchers to establish specificity (see, e.g., Stevens, Hynan, & Allen, 2000) and to meet the criteria for ESTs (see Section IA in the Appendix). Certainly, the clinical scientists supporting various ESTs claim that the efficacy of their treatments is due to the specific therapeutic actions prescribed in the respective manuals. However, specificity, as conceptualized by the medical model, is not supported by the evidence produced by the very clinical trials that are used to identify specific treatments for specific disorders (Wampold, 2001b). First, meta-analysis, the critical analytic tool to synthesize data in the EBP approach, has shown conclusively, from the origins of the procedure by Smith and Glass (1977; Smith, Glass, & Miller, 1980) to contemporary analyses, including treatments that are designed as empirically supported, that no particular treatment or approach is demonstrably superior to another, across disorders or within disorders (Wampold, 2001b; Wampold, Minami, Baskin, & Tierney, 2002; Wampold et al., 1997). The equivalence of outcomes is contrary to what would be expected if a few treatments contained ingredients that were potent while the others used inert procedures. Second, precious little evidence has been found that any treatment works through the mechanism hypothesized. When the cognitive components of CBT for depression are removed, the resulting treatment is as effective as CBT (Jacobson et al., 1996). Indeed, meta-analysis has shown that adding or removing purported critical components of efficacious treatments does not attenuate the benefits (Ahn & Wampold, 2001). Furthermore, when placebo controls are well designed (e.g., have the same number of sessions and well-trained psychologists as the active treatment and psychologists are not proscribed from discussing topics of importance to the patient), the placebo treatments benefit the patients almost as much as the active treatments (Baskin et al., 2003). Additionally, adherence to a protocol does not seem to be related to better outcomes; indeed, strict adherence seems to detract from the alliance and result in poorer outcomes (Wampold, 2001b). Finally, there is little evidence that purported mechanisms either moderate or mediate the relationship between treatment and outcomes as theoretically predicted (Wampold, 2001b).

However, there is persuasive evidence that the factors common to all, or at least most, therapies are responsible for the benefits of such treatments. Psychologist contributions to outcomes overwhelm treatment differences—the person of the psychologist is critical (Wampold, 2001b; see below). The most researched common factor—the alliance between the psychologist and the patient—has been found to be a robust predictor of outcome, even when measured early in therapy (A. O. Horvath & Symonds, 1991;

Martin, Garske, & Davis, 2000). Moreover, the alliance predicts future improvement, even when taking into consideration symptom reduction, so the alliance does not simply reflect past improvement (Klein et al., 2003). There is compelling evidence that it makes more sense to think of elements of the relationship as being empirically supported rather than particular treatments (see Norcross, 2002). Next, we examine how the development of evidence-based treatment movements ignored certain aspects of psychological treatments and emphasized others. It is interesting to note that the omissions we discuss are related to various aspects of therapy that are deemed important by psychologists and patients.

The Omissions

Although the randomized design is a major accomplishment and was instrumental to the development of modern medicine, in many respects, various aspects of the phenomenon being studied were ignored by use of this design (Danziger, 1990; Wampold, 2001a). We focus on two primary omissions: the role of the psychologist in treatments and the patient's subjective experience.

Role of the Psychologist

At the origins, the randomized design was used extensively in three areas: education, agriculture, and medicine. The randomized design was touted by psychology—as that profession sought to show its usefulness to the public and to other professions—as a means to test the effectiveness of educational interventions; the customers were educational administrators who possessed both money and power (Danziger, 1990). The teachers, predominantly low-paid women, were considered unimportant. Thus the randomized design was used to test the relative effectiveness of administrative interventions and the role of the teacher was not considered in the design (see, e.g., McCall, 1923). R. A. Fisher applied his statistical expertise, having been banned from psychometrics by the imperious Karl Pearson, to the field of agriculture, where the emphasis was on soil, fertilizers, strains, and farming. The practitioner (i.e., the farmer) was presumed to be able to apply the practice uniformly and was not considered a source of error. Given the best fertilizers, seeds, rotation practices, and so forth, the farmer (whether Farmer Jean or Farmer Pat) would then be able to optimize production.

Similar to education and agriculture, in medicine, the variance attributable to the physician was considered unimportant, although the situation is more complicated. Recall that medicine was interested in the specific effects of drugs and procedures. Mesmer was discredited not because his treatments were ineffective but because the theoretically postulated animal magnetism was shown to be fallacious. The emphasis on medicine thus is on the demonstration that the specific ingredients are responsible for the benefits, not expectation, hope, remoralization, or any other psychological processes induced by the practitioner. Consequently, the placebo treatment is designed to control for those factors, and the administrator of the treatment is blind to the treatment condition so as not to influence the patient. Clearly, variance attributable to physicians would be uninteresting and unimportant. Indeed, Mesmer was distraught to be informed that Charles Deslon, a former student and assistant, would administer the treatments as Mesmer felt that Deslon was not competent do so; the Royal Commission's and medicine's response was that the treatment per se should be sufficient to benefit the patient through the ingredients and any benefit due to the charisma, warmth, hope, or skill of the physician was irrelevant.

There are two consequential issues caused by ignoring the providers of treatment (Crits-Christoph et al., 1991; Crits-Christoph & Mintz, 1991, Wampold, 1997, 2001a, 2001b; Wampold & Serlin, 2000). The first issue is that ignoring variability among psychologists inflates treatment effects because observed differences between treatments are due, in part, to the variation among psychologists selected for a particular study (see Wampold & Serlin, 2000). That is, every study that ignores the random variability of outcomes among providers, should it exist, increases the likelihood that it will be concluded that one treatment is superior to another and also increases the estimate of the differences between treatments. The other issue is that, over the years, when psychologists' variability has been examined (typically as reanalyses because the primary studies did not consider psychologists in the design), it appears that the variability among providers is far greater than the variability among treatments (Wampold, 2001b). For example, Crits-Christoph and colleagues (Crits-Christoph et al., 1991; Crits-Christoph & Mintz, 1991) have estimated that the variance due to providers ranges upward to 70% on specific measures and averages about 8%, which is considerable given that the variance due to treatments is at most 1% (Wampold, 2001b; Wampold et al., 1997). In the National Institute of Mental Health Collaborative Study for the Treatment of Depression, treatment accounted for 0% of the variability in outcomes, whereas providers, who were selected for their expertise, trained to adhere to the protocols, and supervised, accounted for about 8% of the variance (Kim, Wampold, & Bolt, in press). It should be noted that ignoring provider effects was not due to inadequate statistical theory and procedures; analyses of variance that were sufficient to take into account providers were initially developed in the 1920s and were completed in the 1940s (Serlin, Wampold, & Levin, 2003). Nevertheless, few if any psychotherapy studies have been designed to assess psychologist effects, despite the fact that ignoring them biases the results (Wampold & Serlin, 2000). Thus, clinical researchers have been focusing on the one aspect of therapy that seems to makes little difference (i.e., the type of therapy delivered) while ignoring an important source of variance that does matter—the psychologists (Wampold, 1997, 2001b). Some researchers would even go so far as to restrict their conclusions to the particular psychologists in the clinical trial (i.e., treat psychologists as a fixed factor) so as to increase the power to find treatment differences (cf. Serlin et al., 2003; Siemer & Joormann, 2003).

Ignoring psychologist effects appears to be a legacy of the original usage of randomized designs in education, agriculture, and medicine. However, it is clear that the criteria for ESTs (see the Appendix) as well as the procedures inherent in the EBP movement have continued to perpetuate the omission of psychologists from consideration. The emphasis on treatment, rather than psychologists, is one about which the evidence-based treatment movement is quite proud:

> Psychiatry was one of the first medical specialties to use extensively the randomized controlled trial, and one of the founding principles of the profession of clinical psychology in the 1950s was that practice

> should be based on the results of experimental comparisons of *treatment* [italics added] methods. (Geddes et al., 1997, p. 1484)

Subjective Experience of the Patient

Clinical trials emphasize the effect of treatment on patients, with a focus on symptoms of particular disorders. This involves a shift away from the introspective tradition of Wilhelm Wundt in Germany and Edward Tichenor in the United States, which relied on expert observations of trained psychologists or the participants' own phenomenological experience, and toward the French medical studies and the British measurement tradition, which involved taking measurements of participants. Thus, "what was desired was knowledge of the individuals as *the objects of intervention rather than as the subjects of experience* [italics added]" (Danziger, 1990, p. 67). The experience of the patients were considered unimportant, a tradition that persists in that symptoms and signs of a disorder are paramount in clinical trials in psychology as well as medicine; rarely are patients asked about their experience with treatments and with their disorders. The meaning attributed to one's illness appears to be irrelevant from a "scientific" perspective (Brody, 1997). Psychologists, however, are acutely aware of the importance of the patient's subjective experience and the need to take it into consideration.

Summary and Implications

History is constructed by events and actors making choices; it is not an outcome of manifest destiny. As can be seen from our prospective examination of the history of the evidence-based practice movement, choices have been made and omissions have occurred. The evidence-based treatment movement places emphasis on *treatments* when it has been found that the type of treatment accounts for very little of the variability in outcomes; on the other hand, aspects of treatment that are valued by psychologists and patients and that have been shown to account for variability in outcomes have been ignored.[2] If evidence were taken seriously, one could easily build the case that the attempt to identify particular treatments as privileged is unjustified. The original goals of the EST initiative were laudable: "if clinical psychology is to survive in this heyday of biological psychiatry, APA must act to emphasize the strength of what we have to offer—a variety of psychotherapies of proven efficacy" (Task Force on Promotion and Dissemination of Psychological Procedures, 1995, p. 3). Thus, adopting a medical model was seen as a means to further the goals of a profession, supporting the philosophy of science notion that the history of science is, in part, an outcome of strategic sociopolitical choices and not manifest destiny—evidence does not come without inference (Hacking, 1975; Latour, 1999). It should be understood, then, that any design of a system to identify, and thus privilege, some treatments over others will inevitably advantage some treatments over others. In this case, EST criteria that emphasize the use of manuals, specific treatments for specific disorders, and placebo-type comparison groups are congruent with characteristics of cognitive and behavioral treatments, so it is no surprise that these treatments predominate the list of ESTs.

What is the future of the evidence-based practice movement? Although we cannot foretell the future, psychologists can and should be a potent force in the debate about how various types of evidence will be used to shape the nature of the practice of psychology. History has revealed how diverse disciplines and actors influenced the course of development of the evidence-based practice movements. It is imperative that psychologists make their voices heard and have a primary say in future developments.

[2] The emphasis of this article has been on the historical development of EBP and ESTs. Although space does not permit a discussion here, it should be recognized that there are alternative schemes for using evidence to inform practice, such as those proposed by APA's Divisions 17 (Society of Counseling Psychology; see Wampold, Lichtenberg, & Waehler, 2002) and 29 (Psychotherapy; see Norcross, 2001) that emphasize common factors and broader perspectives of research in lieu of attempts to identify particular treatments.

References

Ahn, H., & Wampold, B. E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology, 48,* 251–257.

Basham, R. B. (1986). Scientific and practical advantages of comparative design in psychotherapy outcome research. *Journal of Consulting and Clinical Psychology, 54,* 88–94.

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,* 973–979.

Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry, 3,* 257–260.

Brody, N. (1980). *Placebos and the philosophy of medicine: Clinical, conceptual, and ethical issues*. Chicago: University of Chicago Press.

Brody, N. (1997). The doctor as therapeutic agent: A placebo effect research agenda. In A. Harrington (Ed.), *The placebo effect: An interdisciplinary exploration* (pp. 77–92). Cambridge, MA: Harvard University Press.

Buranelli, V. (1975). *The wizard from Vienna: Franz Anton Mesmer.* New York: Coward, McCann & Geoghegan.

Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Daiuto, A., et al. (1998). Update on empirically validated therapies, II. *Clinical Psychologist, 51,* 3–16.

Crits-Christoph, P., Baranackie, K., Kurcias, J. S., Carroll, K., Luborsky, L., McLellan, T., et al. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy Research, 1,* 81–91.

Crits-Christoph, P., & Mintz, J. (1991). Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *Journal of Consulting and Clinical Psychology, 59,* 20–26.

Danziger, K. (1990). *Constructing the subject: Historical origins of psychological research*. Cambridge, UK: Cambridge University Press.

Fisher, R. A. (1935). *The design of experiments*. Edinburgh, UK: Oliver & Boyd.

Foucault, M. (1965). *Madness and civilization: A history of insanity in the age of reason* (R. Howard, Trans.). New York: Pantheon Books.

Franklin, B. (1785). *Report of Dr. Benjamin Franklin, and other commissioners, charged by the king of France, with the examination of the animal magnetism as now practiced at Paris.* London: Johnson.

Gallo, D. A., & Finger, S. (2000). The power of a musical instrument: Franklin, the Mozarts, Mesmer, and the glass armonica. *History of Psychology, 3,* 326–343.

Gauld, A. (1992). *A history of hypnotism.* Cambridge, UK: Cambridge University Press.

Geddes, J., Reynolds, S., Streiner, D., & Szatmari, P. (1997). Evidence

based practice in mental health: New journal acknowledges an approach whose time has come. *British Medical Journal, 315,* 1483–1484.

Gehan, E., & Lemak, N. A. (1994). *Statistics in medical research: Developments in clinical trials*. New York: Plenum Medical Book.

Gleick, J. (2003). *Isaac Newton*. New York: Pantheon Books.

Gould, S. J. (1989). The chain of reason vs. the chain of thumbs. *Natural History, 7,* 12–21.

Gould, S. J. (1991). *Bully for Brontosaurus.* New York: Norton.

Hacking, I. (1975). *The emergence of probability*. Cambridge, UK: Cambridge University Press.

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology, 38,* 139–149.

Horvath, P. (1988). Placebos and common factors in two decades of psychotherapy research. *Psychological Bulletin, 104,* 214–225.

Hunt, M. (1997). *How science takes stock: The story of meta-analysis*. New York: Russell Sage Foundation.

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., et al. (1996). A component analysis of cognitive–behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64,* 295–304.

Kim, D. M., Wampold, B. E., & Bolt, D. M. (in press). Therapist effects in psychotherapy: A random effects modeling of the NIMH TDCRP data. *Psychotherapy Research.*

Klein, D. N., Schwartz, J. E., Santiago, N. J., Vivian, D., Vocisano, C., Castonguay, L. G., et al. (2003). Therapeutic alliance in depression treatment: Controlling for prior change and patient characteristics. *Journal of Consulting and Clinical Psychology, 71,* 997–1006.

Latour, B. (1999). *Pandora's hope.* Cambridge, MA: Harvard University Press.

Markowitz, J. C., Klerman, G. L., Clougherty, K. F., Spielman, L. A., Jacobsberg, L. B., Fishman, B., et al. (1995). Individual psychotherapies for depressed HIV-positive patients. *American Journal of Psychiatry, 152,* 1504–1509.

Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 68,* 438–450.

McCall, W. A. (1923). *How to experiment in education*. New York: Macmillan.

Mesmer, F. A. (1980). *Mesmerism: A translation of the original scientific and medical writings of E. A. Mesmer* (G. Bloch, Trans.). Los Altos, CA: William Kaufman. (Original work published 1766)

Norcross, J. C. (2001). Purposes, processes and products of the task force on empirically supported therapy relationships. *Psychotherapy: Theory, Research, Practice, Training, 38,* 345–356.

Norcross, J. C. (Ed.). (2002). *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients*. London: Oxford University Press

O'Leary, K. D., & Borkovec, T. D. (1978). Conceptual, methodological, and ethical problems of placebo groups in psychotherapy research. *American Psychologist, 33,* 821–830.

Pattie, F. A. (1994). *Mesmer and animal magnetism: A chapter in the history of medicine.* Hamilton, NY: Edmonston.

Rosenthal, D., & Frank, J. D. (1956). Psychotherapy and the placebo effect. *Psychological Bulletin, 53,* 294–302.

Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The *Consumer Reports* study. *American Psychologist, 50,* 965–974.

Serlin, R. C., Wampold, B. E., & Levin, J. R. (2003). Should providers of treatment be regarded as a random factor? If it ain't broke, don't "fix" it: A comment on Siemer and Joormann (2003). *Psychological Methods, 8,* 524–534.

Shapiro, A. K., & Shapiro, E. S. (1997). *The powerful placebo: From ancient priest to modern medicine*. Baltimore: Johns Hopkins University Press.

Shepherd. (1993). The placebo: From specificity to the non-specific and back. *Psychological Medicine, 23,* 569–578.

Siemer, M., & Joormann, J. (2003). Power and measures of effect size in analysis variance with fixed versus random nested factors. *Psychological Methods, 8,* 497–517.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32,* 752–760.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The benefits of psychotherapy*. Baltimore: Johns Hopkins University Press.

Stevens, S. E., Hynan, M. T., & Allen, M. (2000). A meta-analysis of common factor and specific treatment effects across domains of the phase model of psychotherapy. *Clinical Psychology: Science and Practice, 7,* 273–290.

Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically-validated psychological treatment: Report and recommendations. *Clinical Psychologist, 48,* 2–23.

Walmsley, D. M. (1967). *Anton Mesmer.* London: Hale.

Wampold, B. E. (1997). Methodological problems in identifying efficacious psychotherapies. *Psychotherapy Research, 7,* 21–43.

Wampold, B. E. (2001a). Contextualizing psychotherapy as a healing practice: Culture, history, and methods. *Applied and Preventive Psychology, 10,* 69–86.

Wampold, B. E. (2001b). *The great psychotherapy debate: Model, methods, and findings*. Mahwah, NJ: Erlbaum.

Wampold, B. E., Lichtenberg, J. W., & Waehler, C. A. (2002). Principles of empirically supported interventions in counseling psychology. *The Counseling Psychologist, 30,* 197–207.

Wampold, B. E., Minami, T., Baskin, T. W., & Tierney, S. C. (2002). A meta-(re)analysis of the effects of cognitive therapy versus "other therapies" for depression. *Journal of Affective Disorders, 68,* 159–165.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "all must have prizes." *Psychological Bulletin, 122,* 203–215.

Wampold, B. E., & Serlin, R. C. (2000). The consequences of ignoring a nested factor on measures of effect size in analysis of variance. *Psychological Methods, 5,* 425–433.

(*Appendix follows*)

Appendix

Criteria for Empirically Supported Treatments

Well-Established Treatments

I. At least two good between-groups design experiments demonstrating efficacy in one or more of the following ways:
   A. Superior (statistically significantly so) to pill or psychological placebo or to another treatment.
   B. Equivalent to an already established treatment in experiments with adequate sample sizes.

or

II. A large series of single-case design experiments ($N > 9$) demonstrating efficacy. These experiments must have:
   A. Used good experimental designs and
   B. Compared the intervention with another treatment as in IA above.

Further Criteria for Both I and II:

III. Experiments must be conducted with treatment manuals.
IV. Characteristics of the client samples must be clearly specified.
V. Effects must have been demonstrated by at least two different investigators or investigating teams.

Probably Efficacious Treatments

I. Two experiments showing the treatment is superior (statistically significantly so) to a waiting list control

or

II. One or more experiments meeting the Well-Established Treatments Criteria IA or IB, III, and IV but not V.

or

III. A small series of single-case design experiments ($N \geq 3$) otherwise meeting well-established treatments.

From "Update on Empirically Validated Therapies II" by D. L. Chambless, M. J. Baker, D. H. Baucom, L. E. Beutler, K. S. Calhoun, A. Daiuto, et al., 1998, *Clinical Psychologist, 51,* p. 4. Copyright 1998 by American Psychological Association Division 12. Adapted with permission.

Received February 2, 2004
Revision received April 14, 2004
Accepted June 25, 2004 ■

Call for Nominations

The Publications and Communications (P&C) Board has opened nominations for the editorships of ***Clinician's Research Digest, Emotion, JEP: Learning, Memory, and Cognition, Professional Psychology: Research and Practice,*** and ***Psychology, Public Policy, and Law*** for the years 2007–2012. Elizabeth M. Altmaier, PhD; Richard J. Davidson, PhD, and Klaus R. Scherer, PhD; Thomas O. Nelson, PhD; Mary Beth Kenkel, PhD; and Jane Goodman-Delahunty, PhD, respectively, are the incumbent editors.

Candidates should be members of APA and should be available to start receiving manuscripts in early 2006 to prepare for issues published in 2007. Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. Self-nominations also are encouraged.

Search chairs have been appointed as follows:

- ***Clinician's Research Digest:*** William C. Howell, PhD
- ***Emotion:*** David C. Funder, PhD
- ***JEP: Learning, Memory, and Cognition:*** Linda P. Spear, PhD, and Peter Ornstein, PhD
- ***Professional Psychology:*** Susan H. McDaniel, PhD, and J. Gilbert Benedict, PhD
- ***Psychology, Public Policy, and Law:*** Mark Appelbaum, PhD, and Gary R. VandenBos, PhD

Candidates should be nominated by accessing APA's EditorQuest site on the Web. Using your Web browser, go to http://editorquest.apa.org. On the Home menu on the left, find Guests. Next, click on the link "Submit a Nomination," enter your nominee's information, and click "Submit."

Prepared statements of one page or less in support of a nominee can also be submitted by e-mail to Karen Sellman, P&C Board Search Liaison, at ksellman@apa.org.

The deadline for accepting nominations is **December 10, 2004,** when reviews will begin.