The Placebo Is Powerful: Estimating Placebo Effects in Medicine and Psychotherapy From Randomized Clinical Trials

Bruce E. Wampold, Takuya Minami, Sandra Callen Tierney, Thomas W. Baskin, and Kuldhir S. Bhati
University of Wisconsin—Madison

The logic of the randomized double-blind placebo control group design is presented, and problems with using the design in psychotherapy are discussed. Placebo effects are estimated by examining clinical trials in medicine and psychotherapy. In medicine, a recent meta-analysis of clinical trials with treatment, placebo, and no treatment arms was conducted (Hróbjartsson & Gøtzsche, 2001), and it was concluded that placebos have small or no effects. A re-analysis of those studies, presented here, shows that when disorders are amenable to placebos and the design is adequate to detect the effects, the placebo effect is robust and approaches the treatment effect. For psychological disorders, particularly depression, it has been shown that pill placebos are nearly as effective as active medications whereas psychotherapies are more effective than psychological placebos. However, it is shown that when properly designed, psychological placebos are as effective as accepted psychotherapies. © 2005 Wiley Periodicals, Inc. J Clin Psychol 61: 835–854, 2005.

Keywords: placebo; psychotherapy; depression; clinical trials; placebo effect

The placebo has had a long and controversial history in medicine and psychotherapy, with regard to its definition, the effects that it produces, and the research designs that are employed to understand it (Shapiro & Shapiro, 1997). The term itself, according to Walach (2003), originated from the Latin psalm verse, "Placebo Domino in regione vivorum" ("I shall please the Lord in the land of the living") and was sung in the Middle Ages as a

Correspondence concerning this article should be addressed to: Bruce E. Wampold, Department of Counseling Psychology, 321 Education Building—1000 Bascom Mall, University of Wisconsin, Madison, WI 53706; e-mail: wampold@education.wisc.edu

JOURNAL OF CLINICAL PSYCHOLOGY, Vol. 61(7), 835–854 (2005)

Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jclp.20129

prayer at the deathbed. Because others were often paid to do the singing, the term *placebo* became associated with a "nearly fraudulent replacement of the real" (Walach, p. 178). When it was recognized in the 18th century that most substances given by doctors to patients were not helpful, the term placebo became associated with any substance that the physician knew (or believed) was not remedial for the disorder but was given only to please or satisfy the patient, thus continuing the connotation of fraud or deceit (Shapiro & Shapiro, 1997). Modern medicine, striving to leave behind a history of healers labeled as charlatans (e.g., Anton Mesmer) whose cures were affected by hope, expectation, or remoralization, sought to demonstrate that the effects of medical treatments were not merely "placebo effects" but were due to the treatment's active ingredients. That is, the active ingredients affected the body through direct, not mind-mediated, physiochemical processes. Nevertheless, modern medicine has accepted the notion that placebos produce effects, to varying degrees, but contends that medical treatments provide benefits over and above what a placebo would provide. Indeed, for decades, Beecher's (1955) estimate that the administration of a placebo leads to significant improvements in approximately one third of cases for which responses are subjective was accepted as truth. However, with the exception of a small cadre of researchers intent on identifying and understanding placebo effects themselves, the predominant thrust of medical and psychotherapy treatment research has been on developing and testing treatments that produce effects beyond what the denigrated placebo can produce.

Our goal in this article is to present an analysis of the placebo concept, from which to understand the logic of research designs using placebo treatments, and to estimate the size of placebo effects in medicine and psychotherapy, taking into account nuances of the placebo effect and methodological considerations. This discussion will necessarily raise several thorny issues that demonstrate the ambiguities which saturate understanding the placebo, particularly as the concept of the placebo is transported from medicine to psychotherapy.

Definitions and Designs

Definitions

A useful approach to understanding placebos is imbedded in formal logic and is represented by Grünbaum's (1981) exposition:

> The therapeutic theory ψ that advocates the use of a particular treatment modality *t* to remedy [disorder] *D* demands the inclusion of certain *characteristic* constituents *F* in any treatment process that ψ authenticates as an application of *t*. Any such process, besides qualifying as an instance of *t* according to ψ, will typically have constituents *C other than* the characteristic ones *F* singled out by ψ. And when asserting that the factors *F* are remedial for *D*, ψ *may* also take cognizance of one or more of the non-characteristic constituents *C*, which I shall denominate as 'incidental.' (p. 159)[1]

To illustrate this definition, consider an antibiotic, a characteristic constituent in the treatment of a bacterial infection according to extant medical theories. The antibiotic is often delivered orally and the pill contains ingredients other than the antibiotic. For example, the brand name antibiotic Ceftin contains the active ingredient (i.e., characteristic constituent *F*) cefuroxime axetil as well as inactive ingredients (i.e., incidental

[1]Grünbaum (1981) uses the term *constituents*, whereas psychotherapy researchers are partial to the term *ingredient*. We use the two terms interchangeably.

constituents) microcrystalline cellulose, sodium croscarmellose, hydrogenated vegetable oil, sodium lauryl sulfate, anhydrous colloidal silica, hypromellose, propylene glycol, titanium dioxide, sodium benzoate, methyl hydroxybenzote, and propyl hydroxybenzoate, which are, according to medical theories, not remedial for bacterial infections and are not offered as a treatment for bacterial infections. These incidental factors are necessary to constitute the pill used to administer the treatment. Grünbaum's logic allows for the possibility that the incidental ingredients affect the patient, and thus it is inappropriate to call these ingredients inactive. For example, a patient may suffer an adverse reaction due to allergies to one of the incidental ingredients. Moreover, the treatment *t* is administered within the context of a therapeutic ritual that may affect the patient independently of the characteristic or incidental constituents of the treatment.

A critical aspect of Grünbaum's definition is that what is considered the characteristic constituent depends on the designation of the theory; that is, characteristic and incidental constituents are defined only by reference to a modal theory. It is the theory that stipulates a particular characteristic constituent and considers effects produced by uncharacteristic constituents to be unimportant, even if such incidental constituents may be of importance to other therapeutic theories. For modern medicine and the physiochemical theory on which it rests, treatments involve interventions that affect the anatomy or physiology of the patient, and these changes purportedly benefit the patient. Any effects produced by other aspects of the treatment, including the hope, expectation, remoralization, therapeutic relationship, or the receipt of an explanatory system are often called placebo effects, with the connotation that such effects are unimportant. Interestingly, if one approaches acupuncture through a Western medicine perspective, as the National Center for Complementary and Alternative Medicine (NCCAM) within the National Institutes of Health does, then the benefits of this treatment procedure, to be considered bona fide, must be due to identifiable physiochemical processes, such as the release of endogenous opioids, rather than processes described in traditional Chinese medicine (e.g., alteration of chi) or through psychological processes (e.g., hope, remoralization, or belief) (Wampold, 2001a).

Thus, what is considered characteristic of one treatment can be considered incidental to another treatment (Critelli & Newman, 1984). For instance, interpretation of the transference is central to psychodynamic approaches to psychotherapy whereas behavioral analysis is central to behavioral approaches and the respective ingredients are incidental to the other approach (Waltz, Addis, Koerner, & Jacobson, 1993). What we call *common factors* can be thought of as a set of aspects that are incidental to the major therapies but that are common to all (or most) therapies.

Grünbaum's approach to defining placebos focuses only on whether certain aspects of the treatment are characteristic or incidental to the therapeutic theory, and is silent with regard to the effects that these ingredients produce as well as the mechanisms that are responsible for the effects. Various models have been suggested to explain observed effects and are contained under several rubrics. Some theories address the patient's expectation, including expectation for change or a change in expectation for subjective responses (e.g., Kirsch, 1997; Price & Fields, 1997). Other theories discuss models of meaning making and the role of symbols in the healing context: Two aspects of the placebo heavily imbedded within culture (e.g., Brody, 1997; Morris, 1997). At a more primitive level, some have argued that the placebo is a stimuli conditioned to elicit benefits in the classical conditioning paradigm (e.g., Ader, 1997, Stewart-Williams & Podd, 2004). The usual attributions made about the causal mechanisms of placebo, such as hope, expectation, remoralization, attribution of meaning, or belief, which are all psychological processes, are central concepts in one or more of these theories.

From Grünbaum's perspective, determining to what degree placebo effects exist is an empirical issue, although as we shall see, each therapeutic theory uses research designs aimed to detect effects produced by the characteristic ingredients, often at the expense of detecting effects produced by the incidental ingredients (viz., placebo effects). That is, typically researchers are uninterested in estimating placebo effects as their focus is on the fertile ground of demonstrating specificity.

Clinical Trials

The history of modern medicine is distinguished by the strong desire to demonstrate physiochemical specificity. Anton Mesmer was discredited not because his cures were ineffective, but because medicine, striving to become a profession, desired to show that treatments occurred through the theorized and specified physiochemical basis and were not due to psychological factors, such as hope or expectation. A Royal Commission appointed by the King of France and chaired by Benjamin Franklin utilized single-blind studies (patients were blind as to which objects were magnetized and which were not) to show that the purported active ingredient (i.e., animal magnetism) was not necessary to obtain the cures demonstrated by Mesmer (Darnton, 1968; McNally, 1999). Anton Mesmer was discredited based not on the efficacy of his cure, but rather because the benefit did not occur as a result of the specific ingredients or through the specified physiochemical processes (Darnton, 1968). Therefore, to avoid the label of charlatan it gradually became required that it be demonstrated that the treatment specific ingredients had a beneficial physiochemical effect. Nevertheless, the development of the randomized double-blind placebo control group design, now required by the Food and Drug Administration (FDA) to approve drugs, took several centuries to evolve to its current state (Shapiro & Shapiro, 1997). Essentially, this design compares the treatment to a placebo in which the placebo is indistinguishable from the treatment in every way except that it does not contain the *active* ingredient (in medical studies, a physiochemical ingredient) and all concerned (patient, treatment administrator, and evaluator) are blinded.

The randomized double-blind placebo control group design is illustrated in Figure 1, where the active treatment and the placebo treatment are the typical groups (or *arms*, as

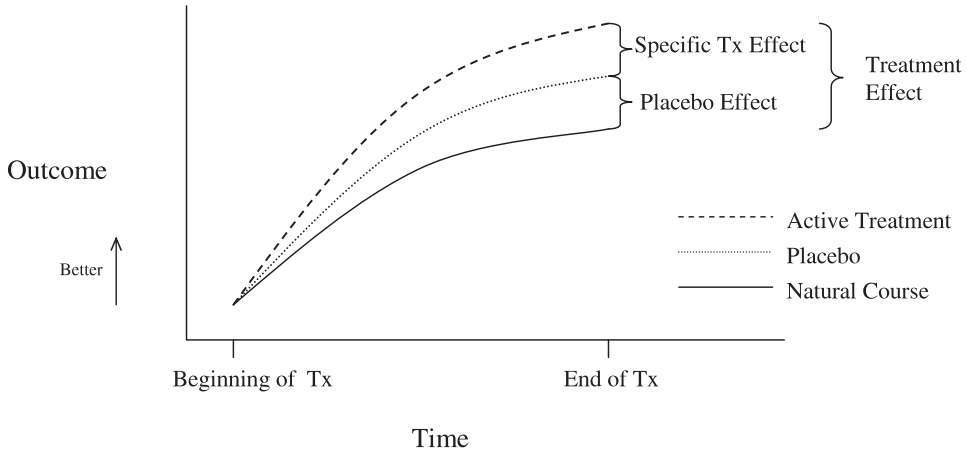

Figure 1. Additive model for placebo effects.

they are often called in medical studies); the natural course of the disorder is portrayed to indicate the various effects of this design. The experiment is designed to test if the specific treatment effect (treatment vis-à-vis placebo) produces a test statistic sufficient to reject the null hypothesis of no differences. The logic is that if the active treatment is shown to be superior to the placebo, which was designed to control for all incidental factors, then the treatment is designated as being effective through the hypothesized physiochemical pathway. When only two groups are used (i.e., treatment and placebo) as is typically the case in medicine, it is impossible to estimate the placebo effect when the natural course of the disorder is unknown, as shown in Figure 1.

There are several assumptions underlying the double-blind randomized placebo control design that defend the design against threats to validity. Of course, there are the usual considerations related to randomization, attrition, and so forth. Critical, however, to the validity of the design is that patients' expectations, hope, and attributions of meaning, which are established through the intertwined concepts of blinding and the ability to distinguish the treatment, be comparable across the treatment and control groups. The treatment and the placebo are indistinguishable if and only if there are no differences in the two treatments that are apparent to the patient or the administrator of the treatment. In a hypothetical experiment, treatment and placebo are indistinguishable if a person receives both and cannot reliably discriminate between the two (in say, a Fisher's lady tasting tea paradigm; see Salsburg, 2001 for an interesting account). Here, indistinguishability is necessary, but not sufficient, for the experiment to be blinded (e.g., the two treatments could be indistinguishable, but the provider of treatment could be knowledgeable of the treatment being provided).

Threats to validity posed by distinguishableness vary, depending on the nature of the treatment and design of the experiment. In any experimental situation in which patients are informed that they will be randomly assigned to conditions, the patients will attempt to determine to which group they have been assigned. If the conditions are truly indistinguishable, then no cues will be provided to the patients and all guesses are random, thereby not affecting the results at the aggregate level. However, a host of factors provides information to patients that may influence their response to treatment. In pharmacological studies, side effects created by the active treatment provide cues that augment the response to the treatment (i.e., subjects recognize by way of side effects that they have been assigned to the treatment group and are receiving the "real" treatment; Greenberg & Fisher, 1997; Kirsch & Sapirstein, 1998), so that although the appearance of the active medication and the placebo are indistinguishable, they are not truly indistinguishable because of their differential side effects. Physical procedures, such as chiropractics, involve certain expectations of physical sensations; the absence of such sensations could threaten validity because patients may use such cues to determine their assignment to groups and the cues may attenuate their expectations for benefits. For example, Sanders, Reinert, Tepe, and Maloney (1990) used a placebo condition that involved only light physical contact, whereas the treatment group received adjustive manipulation at a specific lumbar region. Therefore, it is recommended that in future clinical trials, placebos that produce similar side effects to the active medication should be used to counteract act the potential bias due to "unblinding" (Greenberg & Fisher, 1997; Moncrieff, Wessely, & Hardy, 2004).

The ability to distinguish treatment exacerbates threats to validity when the nature of the treatment and placebo are readily apparent to the provider. If the provider is aware that a placebo is being administered, his or her belief in the intervention will likely produce cues consistent with the attenuated faith in the treatment, whether blatant or subtle, which unavoidably will be communicated to the patient. For example,

psychological treatments involve elaborate procedures lasting over an extended period of time; therapists administering psychotherapy placebos are fully aware that they are providing a sham intervention, and such belief would likely prevent the therapist from faithfully and enthusiastically administering the treatment (Baskin, Tierney, Minami, & Wampold, 2003; O'Leary & Borkovec, 1978), thereby advantaging the treatment over the placebo. Finally, it may be that the placebo is structurally inferior in many respects from the active treatment (e.g., the dose is less), as is often the case with psychotherapy trials, as will be discussed below (see Baskin et al., 2003).

A second set of problems associated with the randomized double-blind placebo control group design is related to the assumption of additivity (Kirsch, 2002a; Kirsch, Scoboria, & Moore, 2002). In Figure 1, the treatment effect is assumed to be the sum of the placebo effect and the specific treatment effect. As Kirsch and colleagues have persuasively argued, it is not logically necessary for the effects of the active treatment to be additive; that is, composed of two components, a placebo effect (equal to the difference between placebo and natural course) and a specific treatment effect (effects due to specific ingredients). The effect of treatment vis-à-vis the natural course of the disorder may be due to the physiochemical properties of the characteristic ingredient entirely or this effect may be larger than the difference between the treatment and placebo outcomes. Consider the case where the treatment and the placebo have equivalent outcomes and the treatment is beneficial because of the physiochemical ingredients and the placebo because of hope, expectation, or remoralization. Indeed, there is evidence that placebos and antidepressants, while exerting nearly equal benefits, exhibit different effects on the brain (Leuchter, Cook, Witte, Morgan, & Abrams, 2002; Mayberg et al., 2002; Olfson, Marcus, Druss, Elinson, Tanielian, & Rincus, 2002). However, there are convincing reasons to believe that specific treatment effects and placebo effects are additive, at least in well-designed studies. First, it is counterintuitive to think that a placebo pill would produce an effect through psychological mechanisms and that the active treatment, indistinguishable from the placebo, would not. Indeed, there is persuasive evidence that the patients' awareness of whether or not they are receiving medical treatment (i.e., open vs. hidden treatments) dramatically affects treatment outcome (Benedetti et al., 2003). Second, there is a strong correlation between placebo effects and treatment effects (Moerman, 2002; Moerman & Jones, 2002; Walach & Maidhof, 1999), which can be understood in the following way. If part of the placebo effect is due to expectations (see Kirsch, 1997), then general knowledge of the effectiveness of treatments of a particular disorder will likely increase patient's expectations. For disorders for which placebos are expected to work, a larger placebo effect would purportedly be obtained. If the effects are additive, a larger treatment effect would be expected as well. If one adopts the additive model, placebo effects will asymptotically approach treatment effects as the degree to which the characteristic ingredients are not remedial for the disorder approaches zero.

The final point to be made regarding randomized double-blind placebo control group design is that typically researchers are attempting to show that the treatment is specific for the disorder. That is, researchers desire to demonstrate that the active treatment is superior to the placebo so that claims can be made about the characteristic ingredients of the treatment to which they have allegiance. In medicine, rarely are no-treatment or waiting list controls used, as the inclusion of such groups would not bolster the case for the specificity, and thus approvability by the FDA, of the medicine or procedure. Moreover, researchers in such trials do not institute procedures that would provide optimal conditions for a placebo effect to occur. For example, in a study of a vaginal antibacterial cream for postpartum infection (Adriaanse et al., 1995), patients who were randomized to active treatment, placebo, or no treatment, were likely unaware of whether the cream was

administered or not, given the myriad of procedures that occur during childbirth; consequently, a placebo effect was unlikely (and of little concern to the researchers).

Despite the limitations of randomized clinical trials for detecting placebo effects, such designs provide useful information for estimating placebo effects and comparing those estimates to treatment effects, providing in some cases estimates of specific treatment effects that are useful for understanding the nature of treatments. In the following section, we will present meta-analytic research on clinical trials in medicine and psychotherapy.

Estimates of Placebo Effects in Medicine and Psychotherapy

Medicine: The Powerful/Powerless Placebo Reconsidered

In 1955, Beecher, based on scant evidence, concluded that the administration of a placebo leads to significant improvement in approximately one third of cases for which responses are subjective; he titled his report *The Powerful Placebo*. Because the focus of medicine at that time was on detecting specific treatment effects, in some sense the size of the placebo effect was irrelevant, creating a disinterest in estimating the placebo effects, particularly in medical clinical trials. However, in 2001, Hróbjartsson and Gøtzsche conducted a meta-analysis of clinical trials published prior to 1999 that compared an active treatment to a placebo treatment and to a no-treatment condition (i.e., all three arms depicted in Figure 1 were included) to estimate the size of the placebo effect. They included all trials of psychological as well as physical disorders and used the primary authors' designation that a placebo was used. Based on these very broad inclusionary criteria, they made the following conclusion:

> We found little evidence in general that placebos had powerful clinical effects. Although placebos had no significant effects on objective or binary outcomes, they had possible small benefits in studies with continuous subjective outcomes and for the treatment of pain. (p. 1594)

However, as we have seen, the clinical trial methodology is not well designed to detect placebo effects and thus, as critics noted, there were many issues with Hróbjartsson and Gøtzsche's study (e.g., Kirsch, 2002b; Kirsch & Scoboria, 2001; Moerman & Jones, 2002; Papakostas & Daras, 2001). First, the Hróbjartsson and Gøtzsche meta-analysis did not consider that placebos are not expected to work uniformly across diseases or disorders (Kirsch & Scoboria, 2001; Shapiro & Shapiro, 1997). "Generally, the presence of anxiety and pain, the involvement of the autonomic nervous system, and the immunobiochemical processes are believed to respond favorably to placebo, whereas hyperacute illnesses (i.e., heart attack), chronic degenerative diseases, or hereditary diseases are expected to resist" (Papakostas & Daras, 2001, pp. 1620–1621). Clearly, there are disorders for which the placebo effect should be large and there are also disorders for which the placebo effect should be nonexistent or small. Aggregating without regard to consideration of heterogeneity of disorders and their amenability to placebo action does not allow for detection of a placebo effect should it exist. Thus, any analysis of the effects of a placebo should differentiate between disorders that are amenable and those that are not amenable to placebo treatment.

Another problematic aspect of the Hróbjartsson and Gøtzsche meta-analysis relates to the lack of attention paid to the mechanism of the placebo. The placebo is a symbol of the healing context and encompasses all aspects of the treatment that have significance for the patient, including the patient's and the physician's beliefs and expectations. Specifically, the patient's expectations regarding the efficacy of interventions are influenced

by multiple cues and are sensitive to the subtle aspects of the healing context, the practitioner, and aspects of the clinical trial. As discussed above, if the patient is unaware that a treatment (or the placebo) has been administered, or if the placebo is distinguishable from the treatment—particularly if it is demonstrably inferior to the treatment—then it would be expected that the placebo effect would be attenuated.

Reanalysis Method. We reanalyzed the studies used in the Hróbjartsson and Gøtzsche study to account for the issues that are raised by estimating placebo effects in clinical trials. Specifically: (a) conditions treated were classified based on their amenability to a placebo treatment; (b) research designs were examined to determine whether the design disadvantaged the placebo treatment or not; (c) the size of the placebo effect was compared to the size of the treatment effect; and (d) placebo effects for subjective and objective measures were compared.

Effect sizes for each study were calculated in the following way. Except as noted below, only the primary measure for each study was analyzed; Hróbjartsson and Gøtzsche provided us with the designated measure and hence, the same measures were used in the original meta-analysis as well as in the present meta-analysis. For each study, three effect sizes were calculated: (a) treatment effect (treatment vs. no treatment), (b) placebo effect (placebo vs. no treatment), and (c) specific treatment effect (treatment vs. placebo). Following the strategy of Hróbjartsson and Gøtzsche, we analyzed studies with continuous outcomes separately from those with dichotomous outcomes.

For each study, two further categorizations were made. First, the degree to which it was expected that a placebo would affect the disorder was determined. Five independent raters (doctoral students in counseling psychology), blind to the results of the study, rated the degree to which the disorder treated in each study could be affected by placebo treatments by classifying each disorder as (a) definitely amenable to psychological factors (e.g., insomnia, chronic pain, depression), (b) possibly amenable to psychological factors (e.g., acute pain, chemotherapy induced nausea, asthma), and (c) not amenable to psychological factors (e.g., anemia, bacterial infection). In all cases, four of the five raters agreed on the classification. It should be noted that amenability to placebo action was operationalized based on the disorder treated, not on the objectivity of the outcome measure. Despite the research that indicates the existence of demonstrable physiological effects attributable to placebos (Leuchter et al., 2002; Mayberg et al., 2002; Olfson et al., 2002), there persists the notion that placebos primarily affect patient self-reports of symptoms (typically labeled *subjective reports*—see Hróbjartsson and Gøtzsche), an assumption that was tested in the present meta-analysis.

Second, whether the design of the study was adequate to estimate the placebo effect or whether research operations attenuated the placebo effect, was evaluated. A study was classified by the five raters as being adequate if all of the following conditions were met: (a) the study was double-blinded, (b) the study participants were aware that they could receive a placebo and were aware when it was administered (i.e., administration was not surreptitious), and (c) the treatment and the placebo were indistinguishable (despite the problem regarding detection via side effects, we considered pill placebos indistinguishable from active pills). The design components were rated separately and the design was determined to be adequate only if each of the components were adequate; agreement on the final determination of adequacy was unanimous based on agreement of four of the five raters. In summary, within the groups of continuous and dichotomous outcomes, studies were classified into six categories by crossing amenability to psychological factors (three levels) by adequacy of research design (two levels). Psychotherapy studies, which will be examined more fully in the next section, were classified as amenable to

placebo, but because such studies are not double-blinded, their study designs were classified as not adequate to estimate the placebo effect.

For the studies with continuous outcomes, standard meta-analytic procedures (Hedges & Olkin, 1985) were used to make two calculations for each of the three effects (viz., treatment, placebo, and relative treatment/placebo effects): (a) an estimate of the effect size d_i for each comparison, and (b) an estimate of the variance of this estimate, (i.e., $\hat{\sigma}_d^2$). For each designated effect, the difference between the means of the two groups was calculated and then divided by the pooled standard deviation of the two groups, and then adjusted to yield an unbiased estimate of the population effect size. To aggregate the effect, we weighted each study's d_i by the inverse of its variance and combined these weighted effect sizes to yield the aggregated effect size estimate $d+$ for each group of studies (Hedges & Olkin, 1985). In addition, the standard error of this estimate $\hat{\sigma}_{d+}^2$ was calculated and used to calculate the 95% confidence interval.

A similar strategy was used for the dichotomous variables. First, an odds ratio, o_i for the two groups being compared was calculated and then transformed to an approximately normal distribution by taking the natural logarithm, (i.e., ln o_i). The variance of the transformed score was calculated and used to construct the 95% confidence intervals. The scores as well as the endpoints of the confidence intervals were then returned to odds ratios by applying the inverse of the natural logarithm (Fleiss, 1994).

For both continuous and dichotomous data, the following hypotheses were made:

1. The placebo effect would be detected (a) when the disorder was amenable to the psychological aspects of the placebo, and (b) when the quality of the research design was adequate.
2. The placebo effect and the treatment effect would not be statistically different (a) when the disorder was amenable to the psychological aspects of the placebo, and (b) when the quality of the research design was adequate.
3. For studies with adequate designs, as the amenability to placebo decreased, the placebo effect detected would also decrease such that the active treatment would become more effective relative to the placebo.
4. There would be no difference in effect sizes of placebos for subjective (patient reported) and objective (data obtained by physiological tests or objective records, but not subjective ratings of evaluators) measures, when both types of measures were used within the same study.

Reanalysis Results. The results for the continuous measures and dichotomous measures are presented in Tables 1 and 2, respectively. Generally, the results were consistent with the hypotheses and thus demonstrated the existence of the placebo effect.

The first hypothesis was that a placebo effect would be detected when the disorder was amenable to placebo action and when the quality of the research design was adequate. For such studies with continuous outcome measures, the placebo effect was statistically larger than zero (viz., $d+ = .29$, $p < .05$). For such studies with dichotomous outcomes, the aggregated odds ratio was not statistically different from 1.00 (viz., $o+ = .99$, *ns*). However, it should be noted that for these studies with dichotomous outcomes, the data did not support the efficacy of the active treatments either ($o+ = .89$, *ns*)

The second hypothesis was that for this same set of studies (i.e., disorders amenable to treatment and adequate design), the placebo effect and the treatment effect would not be statistically different (i.e., there would be no specific treatment effect). For both types of outcome measures, as hypothesized, specific treatment effects were not statistically

Table 1
Effect Sizes for Studies With Continuous Outcome Variables

Amenability and Design	No. Studies[a]	Treatment			Placebo			Specific		
			CI			CI			CI	
		d+	LB	UB	*d*+	LB	UB	*d*+	LB	UB
Definitely										
Adequate	5	.24	.00+	.47	.29	.06	.52	−.05	−.29	.18
Not adequate	29	.83	.69	.97	.23	.08	.37	.58	.44	.71
Possibly										
Adequate	6	.29	.11	.47	.17	−.01	.36	.19	.01	.37
Not adequate	30	.61	.51	.71	.33	.22	.43	.27	.17	.38
No										
Adequate	7	.84	.72	.96	−.03	−.14	.08	.65	.53	.77
Not adequate	2	1.12	.64	1.60	−.11	−.55	.33	1.25	.81	1.70

Note. CI = Confidence Interval; LB = Lower Bound; UB = Upper Bound. Effect sizes are statistically significant if the confidence interval does not include 0.
[a]The total number of studies with continuous variables (= 79) do not equal the original meta-analysis (= 82) because data could not be obtained from two studies and one study did not have an active treatment group.

significant (viz., for continuous variables, $d+ = -.05$, *ns*, and for dichotomous variables $o+ = .89$, *ns*), indicating that placebos produced effects comparable to the treatments.

The third hypothesis was that as the amenability to placebo treatment decreased, the placebo effect detected would also decrease; in addition, the active treatment would become more effective relative to the placebo (i.e., there would be an increasing specific

Table 2
Effect Sizes for Studies With Dichotomous Outcome Variables

Amenability and Design	No. Studies[a]	Treatment			Placebo			Specific		
			CI			CI			CI	
		o+	LB	UB	*o*+	LB	UB	*o*+	LB	UB
Definitely										
Adequate	6	.89	.72	1.09	.99	.81	1.23	.89	.73	1.10
Not adequate	6	.67	.43	1.04	.97	.63	1.50	.73	.47	1.11
Possibly										
Adequate	4	.84	.45	1.57	.96	.52	1.74	.83	.44	1.58
Not adequate	7	.66	.54	.82	.95	.78	1.17	.71	.57	.88
No										
Adequate	4	.85	.57	1.27	.93	.62	1.39	.91	.61	1.37
Not adequate	1	.08	.04	.18	.77	.52	1.14	.11	.05	.22

Note. CI = Confidence Interval; LB = Lower Bound; UB = Upper Bound. Effect sizes are statistically significant if the confidence interval does not include 1.
[a]The total number of studies with dichotomous variables (= 28) do not equal the original meta-analysis (= 32) because data could not be obtained from one study and three studies did not have active treatment groups.

treatment effect). For the studies with continuous measures with adequate research designs, the size of the placebo effect was related to whether the disorder was amenable to psychological factors (viz., placebo effects for definitely, possibly, or not amenable were $d+ = .29$, .17, and $-.03$, respectively), as predicted. Moreover, when the disorder was not amenable to placebo action, the treatment was demonstrably superior to the placebo (i.e., the specific treatment effect was $d+ = .65$). For the studies with dichotomous outcomes with adequate designs, analysis of the trend of the placebo effect was not informative because there was no placebo effect (or treatment effect, for that matter) for any level of amenability to placebo action.

The fourth hypothesis was that within studies that contained the subjective and objective measures, there would be no difference in the size of the placebo effect. In all, 19 studies with continuous data contained both types of measures (there were insufficient dichotomous studies with both types of measures to conduct an analysis). The differences in effect sizes were calculated by subtracting the effect size for the objective measure from the effect size for the subjective measure. The appropriate variance for the difference was then calculated, using the methods for stochastically dependent effect sizes (Gleser & Olkin, 1994, § 4.1, using a correlation between measures of .5; see Wampold et al., 1997). If subjective measures are more sensitive to placebos than objective measures, the effect size for this difference should be statistically greater than zero. In this study, the effect size for this difference was small and not significant ($d = .11$).

Conclusions. When studies used in the Hróbjartsson and Gøtzsche (2001) meta-analysis were disaggregated based on the adequacy of the design and the degree to which the disorder was amenable to psychological factors, evidence for a placebo effect was indeed found. Specifically, for the adequately controlled studies, effects produced by placebo treatments of disorders amenable to psychological factors approached the size of effects produced by treatments.

Hróbjartsson and Gøtzsche concluded that placebos were ineffective except for small but significant effects in studies with continuous subjective outcomes and for the treatment of pain. Contrary to their findings, we found that relatively large effects were produced by adequately conducted studies of disorders amenable to placebo effects when continuous outcomes were used. Effects for objective measures were found to be comparable to those for subjective measures. For the studies with dichotomous outcomes, the ineffectiveness of the active treatments precluded the finding of placebo effects. Thus, it can be seen that in disorders amenable to placebos (i.e., where it is plausible that they could provide an effect via the expectations they create), placebo effects were comparable to treatment effects, thereby establishing the existence of placebo effects.

The conclusions about the power of the placebo vis-à-vis no treatment are limited because there were very few studies ($n = 11$) that employed adequate research designs, involved disorders that theoretically would be amenable to treatment by placebos, and contain both active treatments and no treatment conditions. Moreover, many studies in this data set contained treatments that were not effective or were marginally effective; because the placebo effect theoretically should not exceed treatment effect, the size of the placebo effect in these studies was thereby restricted. It is clear that the clinical trials examined by Hróbjartsson and Gøtzsche were not designed to detect placebo effects, yet when re-aggregated by considering amenability to placebo and adequacy of design, placebo effects are present and their size approaches the size of treatment effects. The result that objective measures produce placebo effects comparable to placebo effects obtained by subjective measures of patients' self-report is an important finding that indicates that the placebo effect is not a superficial phenomenon.

Psychotherapy: A Powerful Placebo?

The results of many studies containing placebo control groups that investigated treatments of psychological disorders, particularly depression, produces an anomaly. The first result is that psychological treatments of depression are generally equivalent to psychopharmacological treatments of depression (Elkin et al., 1989; Hollon et al., 1992). Recent analyses of studies submitted to the FDA for approval of selective serotonin reuptake inhibitor (SSRI) antidepressants produce the second result that indicates that pill placebos are nearly as effective as the SSRI drugs (there was a small, but statistically significant advantage for SSRIs; Kirsch, Moore, Scoboria, & Nicholls, 2002). The final result is that that psychological treatments have been found to be generally superior to psychological placebos. Various meta-analyses estimate the effect size of the psychological treatment/psychological placebo difference to be in the neighborhood of $d = .40–.60$ (Lambert & Ogles, 2004; Stevens, Hynan, & Allen, 2000), which given the robust estimate of treatment effects (viz., treatment vs. no treatment) of .80, is quite large. So, it appears that psychological placebos are not producing the same effects as pill placebos, a result that is unimportant for those focused on treatment efficacy, but vitally important for understanding placebo effects. We attempt to resolve this apparent contradiction through an examination of conceptual and empirical analyses of placebos used in psychotherapy studies.

Over the years, many have commented on the nature of placebo treatments as a control for psychotherapeutic treatments in clinical trials (Basham, 1986; Borkovec & Nau, 1972; Brody, 1980; Horvath, 1988; Lambert & Ogles, 2004; O'Leary & Borkovec, 1978; Sheppard, 1993; Wampold, 1997, 2001a, 2001b). A brief review of issues with psychological placebos as well as analyses of these controls provides insights into not only placebos but the enterprise of psychotherapy as well.

The development of the randomized double-blinded placebo control group design occurred during the late 1940s and early 1950s (Gehan & Lemak, 1994; Shapiro & Shapiro, 1997) and it was shortly thereafter that it was suggested that the logic of the design be used in psychotherapy research:

> 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. (Rosenthal & Frank, 1956, pp. 299–300).

The adaptation of the medical placebo to psychotherapy brings us back to Grünbaum's (1981) conceptualization of incidental ingredients. There are three ways to construct a psychotherapy placebo, if one follows Rosenthal and Frank's (1956) suggestion and uses Grünbaum's conceptualization. However, only the controls developed under the third alternative have traditionally been considered to be placebos.

In the first alternative, one could remove the ingredients that are considered to be remedial according to the psychotherapy approach (e.g., the cognitive-behavioral techniques from cognitive-behavioral treatments) and add ingredients which are incidental to the treatment according to the theory that underlies the psychotherapy approach (e.g., add psychodynamic techniques to the treatment from which the cognitive-behavioral techniques have been removed). Properly reconstructed, the new therapy now represents a treatment intended to be therapeutic (e.g., psychodynamic), but is considered not to be remedial from the perspective of the designated treatment (e.g., cognitive-behavioral). Of course, what results are comparisons of two treatments, both bona fide, each active from

its own perspective and not from the other (Critelli & Neumann, 1984). When treatments intended to be therapeutic are compared, there is persuasive evidence that differences among treatments are small or nonexistent (Lambert, 2004; Wampold, 2001a; Wampold et al., 1997).[2] Thus, when conceptualized in this manner, the placebo appears to be as effective as the treatment.

There is one issue regarding using another therapy as a placebo that is critical. The use of this strategy to construct the placebo places theory at the level of the psychotherapeutic approach. Accordingly, psychodynamic treatment would be a placebo for cognitive-behavior therapy (CBT) provided the theory involves cognitive explanations for the disorder. In medicine, the construction of placebos places the theory at the physiological/anatomical level and does not place explanatory systems at the disorder level. For example, diuretics for the treatment of hypertension would not be used as placebo for beta-blockers because theory is at the level of physiochemical understanding of hypertension and not at a particular explanation for hypertension (i.e., the goal is to control for non-physiochemical causes, viz., psychological causes). Analogously, psychodynamic treatment could be considered an inappropriate placebo for CBT because both are psychological explanations and neither psychodynamic nor cognitive explanations are incidental, using Grünbaum's terminology, at the psychological explanatory level.

The second strategy for constructing a placebo is to remove one or more of the characteristic ingredients without adding anything to the treatment, yielding the dismantling design (Borkovec, 1990), which in many ways meets the criteria of experimental design better than the alternatives. The comparison treatment thus has the ingredients of the treatment save one or a few critical ones, providing a test of whether the removed ingredient or ingredients are necessary to produce the benefits provided by the full treatment. Although the two treatments are distinguishable, they are more similar than any other comparisons found in psychotherapy clinical trials.

A well-known dismantling study was conducted by Jacobson and his colleagues (Jacobson et al., 1996) who compared cognitive therapy, cognitive therapy without identification and modification of core schema, and cognitive therapy without identification and modification of schema and without identification and modification of automatic thoughts. Essentially, they removed one or two critical cognitive components of cognitive therapy, yielding at the end a treatment based primarily on behavioral activation, which at the level of cognitive therapy, is incidental to the theory. Jacobson et al. found that all three groups produced comparable outcomes, suggesting that if the dismantled groups are considered placebos in the sense that they do not have one or more of the characteristic ingredients, then the placebos were as effective as the treatment. Ahn and Wampold (2001) conducted a meta-analysis of dismantling studies and additive studies (where a critical ingredient is added rather than removed) and found that adding or removing ingredients that are theoretically purported to be critical did not affect the outcomes produced. Therefore, when the second strategy for constructing placebos for psychotherapy is used, the placebo treatments are not demonstrably deficient compared to the treatment.

The dismantling design as a means to construct a placebo according to Grünbaum's definition illustrates some issue with placebos in psychotherapy as well as with Grünbaum's conceptualization. Jacobson et al. (1996) found that removing the cognitive components of cognitive therapy did not attenuate the efficacy of the treatment. But it could

---

[2] Of course, when two seemingly different treatments are equally effective, then the proponents of one approach can claim that the other approach implicitly uses their ingredients, as is the case of those who claim that EMDR works because it essentially is an exposure-based treatment (see Lohr, DeMaio, & McGlynn, 2003).

be argued that the dismantled treatment is not a placebo, but rather is a behavioral treatment. However, according to Grünbaum, from the standpoint of cognitive therapy, the behavioral component is incidental and thus can be contained in the placebo. But on the face of it, this is ludicrous because cognitive theorists would give credence to the behavioral components (i.e., they would claim that the behavioral components are not incidental). But the slope is slippery, because these same theorists would recognize that the therapeutic relationship could not be classified as incidental either. Perhaps a better way is to use Waltz et al.'s (1993) classification of such elements as *essential but not unique* rather than use the dichotomy of *characteristic* and *incidental*—the relationship is essential to conducting cognitive therapy but it is not unique (and indeed, is common). However, Waltz et al.'s system leads to its own issues when designing psychotherapy placebos, as will be discussed below.

The third, and typical, strategy for constructing placebos involves using treatments that control for the common factors and go by names such as *alternative treatments*, *supportive counseling*, *non-directive therapy*, *credible attention placebo*, and *common factor control*. These control treatments do not have a cogent theoretical rationale (i.e., they are not treatments intended to be therapeutic) nor do they have therapeutic actions consistent with coherent change principles, but typically involve to varying degrees, a relationship with a trained provider, support, empathic responding, and purported expectation that the treatment will be effective. It is these types of controls that produce benefits that are considerably smaller than the effects produced by active psychotherapies.

The problems with the third strategy are many (see Basham, 1986; Borkovec & Nau, 1972; Brody, 1980; Horvath, 1988; Lambert & Ogles, 2004; O'Leary & Borkovec, 1978; Sheppard, 1993; Wampold, 1997; see Wampold, 2001a and 2001b for a more thorough discussion). The first problem is related to the order of the theory. Placebos are purported to operate through the hope, expectation, remoralization, therapeutic relationship, and other psychological processes. Yet, if the order of theory is at the level of "psychology" then these factors are not incidental to the theory; both specific ingredients (e.g., challenging maladaptive thoughts) and common factors (e.g., hope, remoralization) involve psychological processes. Furthermore, even if aspects of the placebo are considered to be incidental, they are incidental in a way that is different from the way in which lactose is incidental in a pill placebo—in psychotherapy, some of the incidental factors are necessary for the delivery of the treatment and therefore become a perspicuous aspect of the treatment (i.e., they are essential but not unique). The most obvious example is the therapeutic relationship, which is necessary for the delivery of psychotherapy and the quality of which is related to the effectiveness of the active ingredients. Nearly all treatment manuals prominently discuss the relationship with the patient—this is very different from the lactose in pill placebos that is not considered to have any effect on delivery of the active ingredient. Therefore, using Waltz et al.'s (1993) designation of essential but not unique elements raises the question: Should these elements be contained in the placebo or not? If they are contained, how can one ensure that they are given comparably? A relationship in the context of a bona-fide therapy, where there is agreement on goals and tasks (for example, in CBT) is very different from a relationship in the context of an "alternative" treatment where typically a rationale, goals, tasks, and other activities common to therapies intended to be therapeutic do not exist. Indeed, a lack of equivalence in the degree to which common factors are delivered or the quality with which they are delivered can be invoked as an alternative hypothesis to explain various outcomes in psychotherapy research (e.g., see Jacobson, 1991).

A second set of problems is created by the contention that placebos control for the common factors in psychotherapy. It should be noted that "common factors" is an ambiguous

term. The relationship is usually core to any discussion of common factors, but relationship differs by treatment method—a relationship with Ellis within the context of rational emotive therapy would have been very different than a relationship with Rogers in the context of client-centered therapy. Furthermore, common factors include much more than a relationship with a caring professional, and some of these aspects cannot be controlled for by psychotherapy placebos. As discussed by Frank and Frank (1991; Wampold, 2001b), common factors include therapeutic rituals (i.e., a set of procedures) consistent with a convincing rationale based in the healing context and delivered by a healer who believes in the treatment. Thus, common factor placebos do not contain all factors that are expected to be included in treatments minus the specific ingredients because having a set of specific ingredients consistent with the explanation provided either explicitly or implicitly to the client *is* a common factor (see Wampold, 2001b). In this regard, it should be recognized that these alternative treatments, stripped of active ingredients, are not simply reductions that result in an experiential or client-centered therapy; experiential or humanistic practitioners or theorists would reject the notion that these alternative treatments, used to control common factors, are consistent with how they would administer bona-fide humanistic or experiential therapies.

The final set of problems with the common factor placebo is centered on the fact that such placebos are clearly distinguishable from the purportedly active treatment. This ability to distinguish is troublesome because the therapists in studies using these controls are aware that they are delivering a treatment not intended to be therapeutic. Given that allegiance has been shown to be related to outcomes (see Luborsky et al., 1999; Wampold, 2001b) and that most placebo treatments are administered by advocates of the active treatment or are trained by advocates of the active treatment, as well as the difficulty in faithfully and enthusiastically administering a treatment known to the provider to be bogus, most likely attenuates the effectiveness of the placebo treatment.

Essentially, it is difficult to adequately develop a common factor-type control in psychotherapy research. Over the years, attempts have ranged from those that provide convincing, but bogus, rationales (see for example, Borkovec and Costello, 1993, for an excellent attempt) and equal doses of treatment to those without rationales, proscription of therapist actions generally acknowledged to be therapeutic, and decreased doses of treatment. For an example of the latter, consider the comparison of supportive psychotherapy, designed as a placebo control for common factors, to interpersonal psychotherapy for the treatment of depressed HIV patients (Markowitz et al., 1995):

> Supportive psychotherapy, defined as noninterpersonal psychotherapy and noncognitive-behavioral therapy, resembles the client-centered therapy of Rogers,[3] 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 lack of treatment, particularly as delivered by empathic, skillful, experienced, and dedicated therapists. 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 (p. 1505).

This brief discussion of the problems with common factor controls in psychotherapy reveals that such controls have aspects that attenuate their potency vis-à-vis the active

[3] It is unclear that the supportive therapy used in this study would resemble the therapy of Rogers, as client-centered therapy involves more than minimal empathic response.

treatment to which it is being compared. These issues include therapists who know they are delivering a treatment not intended to be therapeutic, no rationales or less convincing rationales, the lack of specific therapeutic actions consistent with a rationale, proscriptions against various therapeutic actions, and smaller doses of treatment for the placebo treatment (Baskin et al., 2003; Wampold, 2001b). Although pill placebos are not without problems (e.g., side effects that provide cues to patients about assignment), they offer better controls than do the psychotherapy placebos that attempt to control for common factors. Consequently, it is not surprising to find that pill placebos produce outcomes that approach that of active medications for depression while psychotherapy placebos produce outcomes that fall short of active psychotherapies.

Recently, Baskin, Tierney, Minami, and Wampold (2003) attempted to estimate the effectiveness of common factor-type placebos vis-à-vis generally accepted treatments by discriminating between those placebos that were structurally equivalent to the active treatment and those that were not. Structurally equivalent placebos had the same number and length of sessions as the active treatment, used the same format (e.g., group, family, individual) as the treatment, used therapists with training comparable to that of therapists of the active treatment, involved treatments that were individualized to the patient, allowed patients to discuss topics logical to the treatment, and did not constrain the conversation to neutral topics. If the placebo did not contain all of these elements, it was classified as not equivalent. Structurally equivalent and placebos that were not equivalent produced different sized effects. Results indicated that comparisons between active treatments and structurally placebos that structurally were not equivalent produced larger effects than comparisons between active treatments and structurally equivalent placebos; moreover, the latter comparison produced negligible effects ($d = .15$), indicating that active treatments were not demonstrably superior to well-designed placebos.

Designing a placebo treatment in psychotherapy is difficult. Nevertheless, using various means to construct placebos, it appears that placebo treatments are nearly as effective as active treatments provided the design of the placebo is adequate; although this result may not apply across all disorders as studies using placebo types controls are not uniformly distributed across disorders.

Conclusions

Two major conclusions follow from our analysis of placebo effects in medicine and psychotherapy. The first conclusion is that the placebo effect is robust. With regard to placebo effects in medicine, when disorders were amenable to placebo treatments and the design of the study was sufficient to detect a placebo effect, the placebo effect was indeed present and approached the size of treatment effects. Moreover, the placebo effect was as strong when it was objectively measured as it was when it was subjectively measured. In psychotherapy, it has been claimed that treatments produce effects that are roughly twice as large as placebo effects (Lambert & Ogles, 2004; Wampold, 2001b). However, when psychotherapy placebos are well designed, the placebo effect approaches the treatment effect, a result consistent with pharmacological treatments of psychological disorders.

The second conclusion is that the notion of placebo in psychotherapy is logically complex. Grünbaum's (1981) definition of *characteristic* and *incidental ingredients* have added rigor to conceptualizing a placebo; nevertheless, as discussed here, problematic issues remain. There are aspects of psychotherapy that do not fit neatly into the characteristic or incidental categories. Moreover, determining whether an ingredient is incidental requires specification of a theory and the order of the theory; that characteristic and incidental aspects in psychotherapy belong to the same class when the order of the theory

is "psychological explanation" further complicates the problem. Apart from these issues, in psychotherapy research additional (although related) problems exist because it is not possible to (a) design a control that is indistinguishable from the active treatment, or (b) blind the study from the perspective of the therapist. The search for specificity in psychotherapy requires elaboration of Grünbaum's logic (e.g., Lohr et al., 2003); a healthy debate exists whether to continue this search or to abandon a medical model of psychotherapy and accept an alternative explanation for the benefits of psychotherapy (see Wampold, 2001b).

Despite the disagreements related to specificity in psychotherapy, the results of clinical trials in psychotherapy and medicine indicate that the placebo is indeed powerful in situations where it would be expected to operate. It is clear that the beneficial aspects produced by medicine and psychotherapy involve factors that are not central to respective modal models or the received view of these endeavors.

References

Ader, R. (1997). The role of conditioning in pharmacotherapy. In A. Harrington (Ed.), The placebo effect: An interdisciplinary exploration (pp. 138–65). Cambridge, MA: Harvard University Press.

Adriaanse, A.H., Kollee, L.A.A., Muytjens, H.L., Nijhuis, J.G., de Haan, A.F.J., & Eskes, T.K.A.B. (1995). Randomized study of vaginal chlorhexidine disinfections during labor to prevent vertical transmission of group B streptococci. European Journal of Obstetrics, Gynecology, and Reproductive Biology, 61, 135–141.

Ahn, H., & Wampold, B.E. (2001). A meta-analysis of component studies: Where is the evidence for the specificity of 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.

Beecher, H.K. (1955). The powerful placebo. Journal of the American Medical Association, 159, 1602–1606.

Benedetti, F., Maggi, G., Lopiano, L., Lanotte, M., Rainero, I., Vighetti, S., et al. (2003). Open versus hidden medical treatments: The patient's knowledge about therapy affects the therapy outcome. Prevention & Treatment, 6, Article 1. Retrieved June 1, 2004, from http://journals.apa.org/prevention/volume6/pre0060001a.html

Borkovec, T.D. (1990). Control groups and comparison groups in psychotherapy outcome research. Washington, DC: U.S. Department of Health and Human Services.

Borkovec, T.D., & Costello, E. (1993). Efficacy of applied relaxation and cognitive-behavioral therapy in the treatment of generalized anxiety disorder. Journal of Consulting and Clinical Psychology, 61, 611–619.

Borkovec, T., & Nau, S.D. (1972). Credibility of analogue therapy rationales. Journal of Behavior Therapy and Experimental Psychiatry, 3(4), 257–260.

Brody, N. (1980). Placebos and the philosophy of medicine: Clinical, conceptual, and ethical issues. Chicago: The 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 interdiscipinary exploration (pp. 77–92). Cambridge, MA: Harvard University Press.

Critelli, J.W., & Neumann, K.F. (1984). The placebo: Conceptual analysis of a construct in transition. American Psychologist, 39, 32–39.

Darnton, R. (1968). Mesmerism and the end of the Englightenment in France. Cambridge, MA: Harvard University Press.

Elkin, I., Shea, T., Watkins, J.T., Imber, S.D., Sotsky, S.M., Collins, J.F., et al. (1989). National Institute of Mental Health treatment of depression collaborative research program: General effectiveness of treatments. Archives of General Psychiatry, 46, 971–982.

Fleiss, J.L. (1994). Measures of effect size for categorical data. In H. Cooper & L.V. Hedges (Eds.), The handbook of research synthesis (pp. 245–260). New York: Russell Sage Foundation.

Frank, J.D., & Frank, J.B. (1991). Persuasion and healing: A comparative study of psychotherapy (3rd ed.). Baltimore: Johns Hopkins University Press.

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

Gleser, L.J., & Olkin, I. (1994). Stochastically dependent effect sizes. In H. Cooper & L.V. Hedges (Eds.), Handbook of research synthesis (pp. 339–355). New York: Russell Sage Foundation.

Greenberg, R.P., & Fisher, S. (1997). Mood-mending medicines: Probing drug, psychotherapy, and placebo solutions. In S. Fisher & R.P. Greenberg (Eds.), From placebo to panacea: Putting psychiatric drugs to the test (pp. 115–172). New York: Wiley.

Grünbaum, A. (1981). The placebo concept. Behaviour Research and Therapy, 19, 157–167.

Hedges, L.V., & Olkin, I. (1985). Statistical methods for meta-analysis. San Diego: Academic Press.

Hollon, S.D., DeRubeis, R.J., Evans, M.D., Wiemer, M.J., Garvey, M.J., Grove, W.M., et al. (1992). Cognitive therapy and pharmacotherapy for depression: Singly and in combination. Archives of General Psychiatry, 49, 774–781.

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

Hróbjartsson, A., & Gøtzsche, P.C. (2001). Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. The New England Journal of Medicine, 344(21), 1594–1602.

Jacobson, N.S. (1991). Behavioral versus insight-oriented martial therapy: Labels can be misleading. Journal of Consulting and Clinical Psychology, 59, 142–145.

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.

Kirsch, I. (1997). Specifying nonspecifics: Psychological mechanisms of placebo effects. In A. Harrington (Ed.), The placebo effect: An interdisciplinary exploration (pp. 166–186). Cambridge, MA: Harvard University Press.

Kirsch, I. (2002a). Are drug and placebo effects in depression additive? Biological Psychiatry, 47, 733–735.

Kirsch, I. (2002b). Yes, there is a placebo effect, but is there a powerful antidepressant drug effect? Prevention & Treatment, 5, Article 22. Retrieved June 1, 2004, from http://journals.apa.org/prevention/volume6/pre0060001a.html

Kirsch, I., Moore, T.J., Scoboria, A., & Nicholls, S.S. (2002). The emperor's new drugs: An analysis of antidepressant medication data submitted to the U.S. Food and Drug Administration. Prevention & Treatment, 5, Article 23. Retrieved June 1, 2004, from http://journals.apa.org/prevention/volume6/pre0060001a.html

Kirsch, I., & Sapirstein, G. (1998). Listening to Prozac but hearing placebo: A meta-analysis of antidepressant medication. Prevention & Treatment, 1, Article 0002a. Retrieved June 1, 2004, from http://journals.apa.org/prevention/volume6/pre0060001a.html

Kirsch, I., & Scoboria, A. (2001). Apples, oranges, and placebos: Heterogeneity in a meta-analysis of placebo effects. Advances in Mind-Body Medicine, 17(4), 307–309.

Kirsch, I., Scoboria, A., & Moore, T.J. (2002). Antidepressants and placebos: Secrets, revelations, and unanswered questions. Prevention & Treatment, 5, Article 33. Retrieved June 1, 2004, from http://journals.apa.org/prevention/volume6/pre0060001a.html

Lambert, M.J., & Ogles, B.M. (2004). The efficacy and effectiveness of psychotherapy. In M.J. Lambert (Ed.), Handbook of psychotherapy and behavior change (5th ed., pp. 139–193). New York: Wiley.

Leuchter, A.F., Cook, I.A., Witte, E.A., Morgan, M., & Abrams, M. (2002). Change in brain function of depressed subjects during treatment with placebo. American Journal of Psychiatry, 159, 122–129.

Lohr, J.M., DeMaio, C., & McGlynn, F.D. (2003). Specific and nonspecific treatment factors in the experimental analysis of behavioral treatment efficacy. Behavior Modification, 27, 322–368.

Luborsky, L., Diguer, L., Seligman, D.A., Rosenthal, R., Krause, E.D., Johnson, S., et al. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. Clinical Psychology: Science and Practice, 6(1), 95–106.

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.

Mayberg, H.S., Silva, J.A., Brannan, S.K., Tekell, J.L., Mahurin, R.K., McGinnis, S., et al. (2002). The functional meuroanatomy of the placebo effect. American Journal of Psychiatry, 159, 728–737.

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

Moerman, D.E. (2002). "The loaves and the fishes": A comment on the emperor's new drugs: An analysis of antidepressant medication data submitted to the U.S. Food and Drug Administration. Prevention & Treatment, 5, Article 29. Retrieved June 1, 2004, from http://journals.apa.org/prevention/volume6/pre0060001a.html

Moerman, D.E., & Jones, W.B. (2002). Deconstructing the placebo effect and finding the meaning response. Annals of Internal Medicine, 136, 471–476.

Moncrieff, J., Wessely, S., & Hardy, R. (2004). Active placebos versus antidepressants for depression. The Cochrane Database of Systematic Reviews 2004, Issue 1, Art. No.: CD003012.pub2.

Morris, D.B. (1997). Placebo, pain, and belief: A biocultural model. In A. Harrington (Ed.), The placebo effect: An interdisciplinary exploration (pp. 187–207). Cambridge, MA: Harvard University Press.

O'Leary, K.D., & Borkovec, T.D. (1978). Conceptual, methodological, and ethical problems of placebo groups in psychotherapy research. American Psychologist, 33(9), 821–830.

Olfson, M., Marcus, S.C., Druss, B., Elinson, L., Tanielian, T., & Rincus, H.A. (2002). National trends in the outpatient treatment of depression. Journal of the American Medical Association, 287, 203–209.

Papakostas, Y.G., & Daras, M.D. (2001). Placebos, placebo effect, and the response to the healing situation: The evolution of a concept. Epilepsia, 42, 1614–1625.

Price, D.D., & Fields, H.L. (1997). The contribution of desire and expectation to placebo analgesia: Implications for new research strategies. In A. Harrington (Ed.), The placebo effect: An interdisciplinary exploration (pp. 117–138). Cambridge, MA: Harvard University Press.

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

Salsburg, D. (2001). The lady tasting tea: How statistics revolutionized science in the twentieth century. New York: Henry Holt.

Sanders, G.E., Reinert, O., Tepe, R., & Maloney, P. (1990). Chiropractic adjustive manipulation on subjects with acute low back pain' visual analog pain scores and plasma–endorphin levels. Journal of Manipulative and Physiological Therapeutics, 13, 391–395.

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

Shepherd, M. (1993). The placebo: From specificity to the non-specific and back. Psychological Medicine, 23(3), 569–578.

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.

Stewart-Williams, S., & Podd, J. (2004). The placebo effect: Dissolving the expectancy versus conditioning debate. Psychological Bulletin, 130, 324–340.

Walach, H. (2003). Placebo and placebo effects—A concise review. Focus on Alternative and Complementary Therapies, 8(2), 178–187.

Walach, H., & Maidhof, C. (1999). Is the placebo effect dependent on time? A meta-analysis. In I. Kirsch (Ed.), How expectancies shape experience (pp. 321–332). Washington, DC: American Psychological Association.

Waltz, J., Addis, M.E., Koerner, K., & Jacobson, N.S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. Journal of Consulting and Clinical Psychology, 61, 620–630.

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., 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.