Journal of Consulting and Clinical Psychology
2003, Vol. 71, No. 6, 973–979

0022-006X/03/$12.00 DOI: 10.1037/0022-006X.71.6.973

Establishing Specificity in Psychotherapy: A Meta-Analysis of Structural Equivalence of Placebo Controls

Thomas W. Baskin, Sandy Callen Tierney, Takuya Minami, and Bruce E. Wampold
University of Wisconsin—Madison

Placebo treatments in psychotherapy cannot adequately control for all common factors, which thereby attenuates their effects vis-à-vis active treatments. In this study, the authors used meta-analytic procedures to test one possible factor contributing to the attenuation of effects: structural inequalities between placebo and active treatments. Structural aspects of the placebo included number and duration of sessions, training of therapist, format of therapy, and restriction of topics. Results indicate that comparisons between active treatments and structurally inequivalent placebos produced larger effects than comparisons between active treatments and structurally equivalent placebos; moreover, the latter comparison produced negligible effects, indicating that active treatments were not demonstrably superior to well-designed placebos.

Psychotherapy treatment outcome studies have used the double-blind randomized placebo control design to rule out the effects of various common factors (Goldfried & Wolfe, 1996). This design was originally developed in the United States and the United Kingdom in the 1930s (Gehan & Lemak, 1994; Shapiro & Shapiro, 1997; Wampold, 2001a) for the purpose of holding constant all factors except the medication's active ingredient. Scientific medicine researchers sought to adapt the concept of randomized clinical trials to establish that the benefits of medications were due to physiochemical properties rather than to patients' expectations, hopes, or other psychological processes. The placebo pill, used in the medical double-blind randomized placebo control design as the typical way of controlling all factors incidental to the treatment, is designed to be indistinguishable from the active medication—in appearance, taste, and smell. In this design, it is necessary that the patient, the administrator of the treatment, and the evaluator be unaware of the patient's treatment condition because the design is intended to rule out psychological factors that are incidental to the purported active ingredient. Clearly, for instance, if the patient were aware that he or she was receiving a pill with no active ingredients, the expectation for improvement would be attenuated.

As noted by Shapiro and Shapiro (1997):

> Gold [who developed the design in the United States] advocated a comparison between "an allegedly potent agent and a blank of such physical properties as to render a distinction between the two impossible except through some pharmacologic potency which may exist . . . [the recommended] double-blind procedure which calls for an investigation in which neither the patient nor the doctor is aware of the identity of the two agents until the results are in and analyzed. This is imperative to avoid the influence of subconscious bias . . ." (Gold, 1954, p. 724). The statement by Gold culminated twenty years of pioneering study of methods with which to reliably and validly evaluate the effectiveness of new drugs. (p. 148)

Shortly after the randomized double-blind placebo control group design was adopted in medicine, Rosenthal and Frank (1956) suggested that the design be used in psychotherapy research to rule out factors that are incidental to ingredients specified by the treatment protocol (i.e., to control for the common factors in therapy):

> 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 standpoint 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)[1]

For example, if cognitive–behavioral therapy (CBT) for depression were compared with an adequate placebo control group and found to produce superior outcomes, these results would support the contention that the purported active ingredients in CBT (e.g., altering core schema and challenging irrational thoughts) were responsible for the benefits of the treatment. This assertion could

Thomas W. Baskin, Sandy Callen Tierney, Takuya Minami, and Bruce E. Wampold, Department of Counseling Psychology, University of Wisconsin—Madison.

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, Madison, Wisconsin 53706. E-mail: wampold@education.wisc.edu

[1] Currently, it is not popular to call alternative treatments *placebos* because of the connotations of deception and charade. Consequently, such groups are labeled as *supportive therapy*, *nondirective therapy*, *common factor control*, *credible attention placebo*, and *modest contact*. However, because their purpose is to rule out common factors, they are used in an attempt to emulate the role placebos play in the medical model of testing drug efficacy, and thus, they are called placebo controls herein. At times, the actions of the therapists in these control groups appear to have a Rogerian flavor; however, the "Rogerian" treatment provided would not meet the definition of a bona fide treatment used in this research, nor would such treatment be accepted as a viable experiential therapy as currently conceptualized (see Wampold, 1997, 2001b).

only be made as long as the placebo treatment was sufficiently well designed to rule out factors such as the therapeutic relationship, expectation for benefits, the healing context, and so forth.

On the basis of this rationale, researchers over the years have used the double-blind, randomized, placebo control design to establish the efficacy of specific treatment ingredients in psychotherapy studies by purportedly controlling for the treatment effects due to factors extraneous to the specific treatment. To appropriately use this design to establish the efficacy of a specific psychological ingredient, researchers must meet several conditions. First, all factors but those purported to be the active ingredients of that treatment must be made equivalent (i.e., held constant). That is, the placebo control treatment must be structurally indistinguishable from the specific treatment (e.g., same number of sessions and same treatment modality). In addition, the participants, study administrators, and study examiners must all be blind to the treatment conditions. Finally, other psychological factors, such as participants' expectations and belief in treatment and the credibility of the treatment, need to be controlled (i.e., made comparable for the active and placebo treatments).

Unfortunately, aspects of placebo control groups in psychotherapy attenuate their validity to establish that the benefits of psychotherapy are derived from the specific ingredients (Basham, 1986; Borkovec & Nau, 1972; Brody, 1980; Horvath, 1988; O'Leary & Borkovec, 1978; Shepherd, 1993; Wampold, 1997, 2001a, 2001b). The problems are twofold. First, psychotherapy studies cannot be blind in the manner of placebo controlled medical studies. Quite obviously, the therapist must be aware of the treatment being delivered to follow the treatment protocol (O'Leary & Borkovec, 1978; Seligman, 1995; Wampold, 2001b). It is quite likely that in the absence of therapist blinding, the therapist, knowing that he or she is administering a placebo treatment not intended to be therapeutic, administers the placebo treatment less faithfully and passionately: "It is unlikely that therapists would accept such a placebo if they had to implement it for 20 sessions" (O'Leary & Borkovec, 1978, p. 823).

The second problem is that the placebo is distinguishable from the active treatment; that is, an objective observer could identify the differences. The apparent differences between placebo and active treatment play havoc with the logic of the design. If the study participants are informed that they will be randomly assigned to an active treatment or to a placebo treatment without active ingredients, the apparent differences between the groups would allow the participants to determine that they were assigned to the less desirable treatment. Therefore, the quality of the design is increased by informing the participants that both treatments are equally efficacious, thereby creating a deception (O'Leary & Borkovec, 1978). However, even with the deception, the differences create, to varying degrees, inequivalence in the common factors that are being controlled. For example, the credibility of the rationale of treatment is a potentially powerful healing aspect of treatment (Wampold, 2001a, 2001b); by necessity, the rationales for the treatments differ and, consequently, the credibility of the treatment and the resulting expectations of the participants might differ as well. Indeed, Borkovec and Nau (1972) found that the rationales for control conditions were rated as less credible than the rationales for various active conditions.

Although it is impossible to design a placebo psychotherapy that is indistinguishable from the active treatment, obviously, the more similar the placebo is to the active treatment, the better. At the very least, the two treatments should be structurally equivalent. At the most basic level, structural equivalence involves the dose of treatment, particularly because dose is related to outcome (Howard, Kopta, Krause, & Orlinsky, 1986). Thus, treatment and placebo should provide participants with equal numbers of sessions of equal length. As well, the treatments should use the same modality (e.g., individual or group) and therapists with generally equivalent skill and training. Finally, the placebo treatment should not proscribe therapists behaviors to the extent that the treatment no longer resembles psychotherapy as typically defined; that is, the placebo treatment should allow participants to discuss their particular problem (i.e., they should not have received a standard psychoeducational treatment or they should not have been restricted from discussing their particular issues).

Adequacy of placebo controls is notoriously difficult to examine empirically. One strategy is to rely on participant ratings of credibility and expectation (e.g., Barker, Funk, & Houston, 1988; Bowers & Clum, 1988; Stevens, Hynan, & Allen, 2000). Unfortunately, such ratings are reactive in that treatment progress affects the ratings (O'Leary & Borkovec, 1978); on the other hand, administration of credibility and expectation scales prior to the beginning of treatment results in responses that are based on negligible amounts of information and are therefore unreliable. Moreover, most studies using placebos do not assess participants' expectations.

There have been previous attempts to relate adequacy of placebo controls to outcomes (viz., Barker, Funk, & Houston, 1988; Bowers & Clum, 1988; Stevens et al., 2000). Barker et al. (1988) examined only studies for which the differences in ratings of expectation prior to treatment for the treatment and control groups were not statistically different. Such a procedure is based on participants' judgment of the therapy prior to actually receiving therapy, thereby relying on a small set of cues provided by the researcher; moreover, only 17 studies out of over 200 reviewed provided adequate data to conduct the meta-analysis. Bowers and Clum (1988) as well as Stevens et al. (2000) examined credibility of treatments. Unfortunately, both studies used either participant ratings of credibility subsequent to treatment or, if such ratings were nonexistent, characteristics of the placebo control group, which were considered proxies for credibility. Moreover, pill placebos were considered as credible controls for the active psychological treatment. Surprisingly, in these previous meta-analyses, credibility did not appear to be related to treatment versus control group outcomes, bolstering the conjecture that use of a combination of ratings and proxies for credibility as well as inclusion of pill placebos obscured finding the expected relationship.

The purpose of the present study was to examine the effects related to the structural equivalence of the placebo vis-à-vis the active treatment. It would seem reasonable that research designs using placebo controls would include placebo controls that were structurally equivalent to the treatments in order to make a valid conclusion about the specificity of active ingredients, but surprisingly, as is seen, this is not always the case. An example of structural inequivalence is a comparison of supportive psychotherapy, designed as a placebo control for common factors, with 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, 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 therapists. Sixteen 50-min sessions of interpersonal therapy were scheduled within a 17-week period. The supportive psychotherapy condition had between 8 and 16 sessions, determined by patient need, of 30–50-min duration. (p. 1505)

We hypothesized that degrading the structure of the placebo treatment vis-à-vis the active treatment would attenuate the efficacy of the control, thereby producing larger treatment-placebo differences than better designed placebos. That is, the effect size of studies comparing inequivalent treatments would be greater than the effect size of studies comparing structurally equivalent treatments ($d_{\text{Inequivalent}} > d_{\text{Equivalent}}$).

Method

Study Inclusion Criteria

Studies comparing a psychological treatment (i.e., a specific treatment) intended to be therapeutic with an alternative treatment, placebo treatment, or supportive counseling treatment (labeled placebo control for the purposes of this study) that lacked the active ingredients of the specific treatment were included in this meta-analysis. In addition, the following criteria had to be met: the participants were adults, they were randomly assigned to either a treatment group (i.e., bona fide treatment) or a placebo treatment control group, and they received more than one session of a given treatment. Using criteria similar to those identified in Wampold et al. (1997), we deemed that a psychotherapeutic treatment was bona fide on the basis of the following criteria: the treatment must have been delivered by a trained therapist; the treatment must have involved face-to-face contact and interaction with the patient as the vehicle through which the therapeutic components were delivered (e.g., a videotaped lecture series on mental health issues, played to successive participants, would not be considered a bona fide psychotherapeutic treatment); and the treatment must also have included two of the three following components: (a) the treatment was based on psychological principles (e.g., operant conditioning), (b) the treatment was offered to the psychological community as a viable treatment (e.g., manuals were available), and (c) the treatment contained specific components intended to be therapeutic (e.g., cognitive restructuring). Therefore, a treatment intervention that did not meet all of these criteria was not considered to be bona fide and the study was excluded. A placebo treatment was defined as an experimental condition used in an attempt to control for psychotherapy's common factors by providing a treatment devoid of specific ingredients.

Study Acquisition

Studies meeting the previous criteria were acquired in two ways. First, following the procedures used in Wampold et al. (1997), we made an exhaustive search of six major journals published between 1994 and 2000, inclusive. The following journals were selected because they contained the majority of published psychological outcome studies: *Journal of Consulting and Clinical Psychology*, *Journal of Counseling Psychology*, *Behavior Therapy*, *Behavior Research and Therapy*, *Archives of General Psychiatry*, and *Cognitive Therapy and Research*. The years 1994 to 2000 were selected to ensure that conclusions were germane to current research practices, as the use of placebo controls has evolved over the years (cf. Stevens et al., 2000). Second, an electronic database search was conducted. A query was made of the PsycINFO database using the years 1994–2000 and articles in the English language as the primary criteria. Adding the terms *supportive therapy*, *supportive*, and/or *nondirective* yielded abstracts of 100 articles. These articles were then thoroughly examined by a team of four investigators (i.e., the authors) to determine whether they met the criteria. A total of 27 comparisons of treatments to placebos (derived from 26 studies) meeting the inclusion criteria were identified using these two procedures. However, only 21 of the 27 comparisons included the data necessary to conduct a meta-analysis. Therefore, analyses and results are presented for the final sample of 21.

The treatments used in these studies varied considerably. Examples of the types of bona fide specific treatments included: cognitive therapy or CBT (14 studies), behavioral therapy (2 studies), eye movement desensitization and reprocessing (1 study), coping skills training (2 studies), social skills training (1 study), and interpersonal psychotherapy (1 study). The placebo control treatment groups were labeled supportive therapy (13 studies), credible attention placebo (1 study), relaxation training control (1 study), discussion group (1 study), awareness through movement (1 study), befriending (1 study), nonprescriptive treatment (1 study), cognitive analytical treatment (1 study), and modest contact (1 study).

Structural Rating

We carefully examined the structural qualities of the qualifying studies to test adequately our hypothesis that structural differences in the specific treatment and the placebo treatment would affect the treatment outcomes. Two evaluators (doctoral students in counseling psychology), who were blind to the study hypotheses and not connected with the study in any other capacity (i.e., not the authors), were provided with only the Method sections of all 21 studies. They were instructed to rate each treatment group (i.e., the specific and placebo) on the following components: (a) the number of sessions, (b) the length of sessions, (c) the format (i.e., group or individual), (d) the training of the therapists, (e) whether interventions were individualized to the clients, and (f) whether clients could discuss topics logical to the treatment (e.g., were trauma victims allowed to talk about their trauma?) or whether they were restricted to neutral topics. If the treatments were identified as differing on one or more of these dimensions, they were considered to be structurally different and were labeled as having an inequivalent structure. If they did not differ on any of these dimensions, they were categorized as having an equivalent structure. The raters agreed on the categorization of 16 of 21 studies. The authors resolved disagreements consensually by reviewing the Method sections of the articles. All ratings were made blind to the results of the individual studies. This classification method resulted in 8 studies being classified as having inequivalent structure and 13 studies being classified as having equivalent structure. To rule out the threat that cognitive and/or behavioral treatments might have been unequally represented in the two conditions, we checked the distribution of these treatments and found that they were nearly equally distributed: 6 of 8 and 10 of 13 in the inequivalent and equivalent conditions, respectively, were cognitive and/or behavioral (Fisher's exact test, $p = 1.00$).

Meta-Analysis

A meta-analysis was conducted to test the hypothesis pertaining to structural equivalence. Following standard meta-analytic procedures (Hedges & Olkin, 1985), two calculations were made for each study: (a) an estimate of the effect size (d_i) for the bona fide treatment versus the placebo control and (b) an estimate of the variance of this estimate ($\hat{\sigma}_d^2$).

The estimate of the effect size for each study was calculated through a multistage process. For each dependent variable within a study, the difference between the means of the two treatments was calculated and then

divided by the pooled standard deviation of the two treatments. This value was adjusted to yield an unbiased estimate of the population effect size. These unbiased estimates were then aggregated across all dependent measures within a given study using the method developed by Wampold et al. (1997), resulting in one estimate per study and the standard error of that estimate.

Next, we aggregated the effect sizes for the studies in the inequivalent structure group and the equivalent structure groups, respectively. To do this, we followed Hedges and Olkin's (1985) procedures for weighting each study's effect size by the inverse of its variance and combined these weighted effect sizes to yield the aggregated effect size estimate ($d+$) for each structural group. This procedure gives studies with larger sample sizes more weight. In addition, the standard error of this estimate was calculated according to the methods developed by Hedges and Olkin (1985). Using the standard error of estimate, the 95% confidence interval for the true value of the population effect size was calculated. The effect size for the inequivalent structure was compared with the effect size for the equivalent structure, and it was hypothesized that the former would be greater than the latter. Finally, the homogeneity of the effects within each group (viz., the inequivalent and the equivalent) were examined to determine whether there were additional moderators of effect size not considered in this study.

Results

The results of the meta-analysis of the 21 studies containing 913 participants, with regard to structural equivalence, are summarized in Table 1. The aggregated effect size for the 8 studies identified as having inequivalent structures was .465 (361 participants), an effect size that is similar to that found for the difference between bona fide and placebo treatments (Lambert & Bergin, 1994) and that can be classified as a moderate-sized effect (Cohen, 1988). The 95% confidence interval was .309 to .621, indicating that the population effect size is significantly greater than zero. Moreover, the Q statistic used to index heterogeneity was not sufficiently large for us to reject the null hypothesis that these effects are homogeneous ($Q = 7.87$, compared with a chi-square with $k - 1 = 7$ *df*, in which k is the number of studies; see Hedges & Olkin, 1985).

The equivalent structure group's ($n = 13$ studies, 552 participants) aggregated effect size was .149, with a 95% confidence interval of .005 to .292. Although the effect size comparing active treatment with placebos with equivalent structures was significantly larger than zero, the effect size was extremely small (Cohen, 1988; cf. Wampold, 2001b). The effects for this group were heterogeneous ($Q = 42.81$ compared with a chi-square with $k - 1 = 12$ *df*, in which k is the number of studies; see Hedges & Olkin, 1985).

The critical test of the hypothesis is the comparison of the effect size for inequivalent placebos (viz., $d = .465$) with the effect size for the equivalent placebos ($d = .149$). Using the between-groups test developed by Hedges and Olkin (1985), we compared the test statistic Q_B to a chi-square statistic with one degree of freedom. The obtained value of Q_B was 8.57, which is sufficiently large to reject the null hypothesis that the effect sizes are equal ($p = .003$) in favor of the alternative hypothesis that inequivalent placebos produced larger sized effects than those produced by equivalent placebos.

Discussion

In medicine, randomized double-blind placebo control groups are used to establish the specificity of treatments. The validity of the design in psychotherapy is arguable, although many researchers continue to place their faith in such designs: "The task of demonstrating that specific techniques in complete treatments exceed common factor effects *requires* outcome studies that compare complete treatments with common factor controls" (emphasis added, Stevens et al., 2000, p. 273). However, as discussed in the introduction, placebo controls in psychotherapy suffer from a number of flaws. Thus, it is necessary to examine aspects of the design of the studies (viz., structural aspects of the placebo treatments) to understand the results of placebo controlled research in psychotherapy. In the present meta-analysis, we classified placebo controls as inequivalent and equivalent to allow an examination of factors that are related to the efficacy of placebos vis-à-vis treatments.

Although previous attempts to summarize research involving placebos have examined the credibility of treatments (e.g., Bowers & Clum, 1988; Stevens et al., 2000), the present meta-analysis was the first to examine structural inequalities as an indicator of the sufficiency of placebo controls in psychotherapy. Not surprisingly, psychotherapy treatments produced better outcomes than placebos that were structurally inferior. Indeed, this meta-analysis confirmed that the effect size produced by comparisons to inequivalent placebos (i.e., $d = 0.465$) was approximately equal to that found in various other meta-analyses (cf. Lambert & Bergin, 1994; Stevens et al., 2000; Wampold, 2001b). However, when placebo controls are better designed, the present meta-analysis found that the benefits produced by such treatments were not substantially different from the active treatments to which they were compared (i.e., $d = 0.149$, which although statistically different from zero is negligible). It should be noted, in addition, that the latter effect size is heterogeneous, indicating an unobserved moderator (i.e., the variability among the structurally equivalent controls exceeded that expected, suggesting that some aspect of these comparisons influenced the size of the effect).

The results of this meta-analysis have considerable implications for psychotherapy research and theory. With regard to research, it is clear that the design of the placebo control has an effect on the outcome. The placebos in the inequivalent condition were struc-

Table 1
Effect Sizes and Meta-Analytic Summary Statistics and Tests for Inequivalent and Equivalent Structured Placebos

Placebo group	No. of studies	n	d	$\hat{\sigma}_d^2$	95% CI	Homogeneity Q	p
Inequivalent structure	8	361	0.465	0.0063	0.309, 0.621	7.87	.343
Equivalent structure	13	552	0.149	0.0053	0.055, 0.292	42.81	.000

Note. Between-groups statistic $Q_B = 8.57$, $p = .0034$. CI = confidence interval.

turally impoverished relative to the active treatment, particularly with regard to the amount of contact time between the therapist and the patient. Given that dose of therapy and benefits are related (Howard et al., 1986), the poor outcomes produced by the inequivalent condition were expected. At the very least, researchers need to yoke the structure of the placebo to the structure of the treatment.

The results of previous meta-analyses of placebo controlled research have led to very different conclusions about the relative efficacy of common factors and specific ingredients. Stevens et al. (2000) concluded that "the specific components of psychotherapy exert a beneficial influence over and above the common factors delivered" (p. 286) on the basis of the facts that treatments outperformed placebo controls and that credibility did not appear to moderate the effect size. On the other hand, Wampold (2001b) argued that because placebo controls are necessarily flawed in psychotherapy (i.e., they contain some but not all of the common factors; see Wampold, 2001b, Chapter 5), the superiority of bona fide treatments vis-à-vis placebos is *not* sufficient evidence to establish specific effects. The present study attempted to resolve this interpretive ambiguity empirically. Placebo controls that were structurally equivalent to the treatment produced effects that were nearly equal to those produced by active treatments, whereas placebo controls that were structurally inferior produced demonstrably poorer outcomes than active treatments. Two previous meta-analyses have found that credibility was unrelated to outcome (Bowers & Clum, 1988; Stevens et al., 2000); in those studies, elements of structure were incorporated in the operational definition of credibility and included pill placebos as well.[2] It appears, however, that the structure of comparison groups, objectively determined, has a demonstrable effect on the results. The present meta-analysis shows that the evidence collected from studies that used placebo controls is insufficient to conclude that the specific ingredients are responsible for the benefits of psychotherapeutic treatments.

It should be recognized that all placebos in psychotherapy contain one fundamental flaw: In the placebo condition therapists are aware that they are delivering a treatment not containing actions that are theoretically prescribed by any known psychotherapy treatment and thus might not deliver the treatment with the required allegiance to that treatment. That allegiance is an important determinant of outcome in psychotherapy (Wampold, 2001b) suggests that the therapists may be less enthusiastic, less hopeful, less engaged, or less empathic in such conditions; that is, the common factors in such controls are degraded (O'Leary & Borkovec, 1978; Wampold, 2001b). There are other flaws in psychotherapy placebos as well, but nevertheless, this one flaw precludes ever concluding that the superiority of a treatment to a placebo is unambiguous evidence for specific effects (as opposed to the greater certainty that derives from such designs in medicine, in which the double-blind is possible). Thus, a perspicuous case could be made that the placebo design is inadequate to rule out common factors in psychotherapy. The results of the present meta-analysis suggest that in spite of the flaws, well-designed placebos are nearly as beneficial as active treatments.

Of course, placebo control groups are not the only means to establish that the specific ingredients in treatments produce the benefits of psychotherapy (Borkovec & Miranda, 1999; Wampold, 2001b). When placebos were introduced to psychotherapy researchers in 1956, Rosenthal and Frank emphasized the necessity of showing that the intended psychological mechanism was operative. For example, changes in cognitions and schemas should mediate the administration of CBT and the benefits produced (see Wampold, 2001b, Chapter 5). As well, the dismantling design, in which one (or a few) critical ingredients are removed to ascertain whether the benefits can be attributed to that (or those) ingredients, provides a two-group comparison in which the groups are more similar than is the case in the treatment–placebo comparison:

> One crucial feature of the [dismantling] design is that more factors are ordinarily common among the various comparison conditions. In addition to representing equally the potential impact of history, maturation, and so on and the impact of nonspecific factors, a procedural component is held constant between the total package and the control condition containing only that particular element. Such a design approximates more closely the experimental ideal of holding everything but one element constant . . . . Therapists will usually have greater confidence in, and less hesitancy to administer, a component condition than a pure nonspecific condition. They will also be equivalently trained and have equal experience in the elements relative to the combination of elements in the total package . . . . At the theoretical level, such outcomes tell what elements of procedure are most actively involved in the change process . . . . At the applied level, determination of elements that do not contribute to outcome allows therapists to dispense with their use in therapy. (Borkovec, 1990, pp. 56–57)

The purpose of the study needs to be taken into consideration when designing comparison groups in psychotherapy. If research is focused primarily on determining the efficacy of a treatment rather than on theoretically determining specific ingredients, no treatment controls, treatment as usual, or previously established treatments provide useful comparisons. On the other hand, dismantling designs are particularly suited to determining whether a particular ingredient is critical to the benefits of the treatment. Placebo-type controls appear to be located somewhere between pragmatic designs (i.e., those focused on efficacy) and theoretical designs (i.e., those focused on specific components). In spite of this ambiguous position, the present results, derived from placebo controlled research, suggest that specific ingredients in psychotherapeutic treatments are not responsible for the treatment benefits because treatments that are structurally equivalent to active treatments, but which contain no theoretically prescribed ingredients, produce effects similar to those of active treatments.

[2] The present meta-analysis and that of Stevens et al. (2000) had no studies in common. The primary reason for this is that Stevens et al. included only studies that also contained a wait-list control, and 18 of the 21 included in ours did not. Interestingly, Stevens et al. included no studies published after 1993, whereas we restricted our search to studies published after 1993.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Azrin, N. H., McMahon, P. T., Donohue, B., Besalel, V. A., Lapinski, K. J., Kogan, E. S., et al. (1994). Behavior therapy for drug abuse: A

controlled treatment outcome study. *Behavior Research Therapy, 32,* 857–866.

Barker, S. L., Funk, S. C., & Houston, B. K. (1988). Psychological treatment versus nonspecific factors: A meta-analysis of conditions that engender comparable expectations for improvement. *Clinical Psychology Review, 8,* 579–594.

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

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., & Miranda, J. (1999). Between-group psychotherapy outcome research and basic science. *Journal of Clinical Psychology, 55,* 147–158.

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

Bowers, T. G., & Clum, G. A. (1988). Relative contributions of specific and nonspecific treatment effects: Meta-analysis of placebo-controlled behavior therapy research. *Psychological Bulletin, 103,* 315–323.

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

*Brown, R. A., Evans, D. M., Miller, I. W., Burgess, E. S., & Mueller, T. I. (1997). Cognitive–behavioral treatment for depression in alcoholism. *Journal of Consulting and Clinical Psychology, 65,* 715–726.

*Bryant, R. A., Harvey, A. G., Dang, S. T., Sackville, T., & Basten, C. (1998). Treatment of acute stress disorder: A comparison of cognitive–behavioral therapy and supportive counseling. *Journal of Consulting and Clinical Psychology, 66,* 862–866.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

*Cottraux, J., Note, I., Albuisson, E., Yao, S. N., Note, B., Mollard, E., et al. (2000). Cognitive behavior therapy versus supportive therapy in social phobia: A randomized controlled trial. *Psychotherapy and Psychosomatics, 69,* 137–146.

*Craske, M. G., Maidenberg, E., & Bystritsky, A. (1995). Brief cognitive–behavioral versus nondirective therapy for panic disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 26,* 113–120.

*Edelman, S., Bell, D. R., & Kidman, A. D. (1999). Group CBT versus supportive therapy with patients who have primary breast cancer. *Journal of Cognitive Psychotherapy, 13,* 189–202.

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

Goldfried, M. R., & Wolfe, B. E. (1996). Psychotherapy practice and research: Repairing a strained alliance. *Journal of Consulting and Clinical Psychology, 51,* 1007–1016.

*Goldstein, A. J., de Beurs, E., Chambless, D. L., & Wilson, K. A. (2000). EMDR for panic disorder with agoraphobia: Comparison with waiting list and credible attention-placebo control conditions. *Journal of Consulting and Clinical Psychology, 68,* 947–956.

*Grant, J. R., & Cash, T. F. (1995). Cognitive–behavioral body image therapy: Comparative efficacy of group and modest-contact treatments. *Behavior Therapy, 26,* 69–84.

*Hayes, R. L., Halford, W. K., & Varghese, F. T. (1995). Social skills training with chronic schizophrenic patients: Effects of negative symptoms and community functioning. *Behavior Therapy, 26,* 433–449.

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

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

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist, 41,* 159–164.

*Kirkby, R. J. (1994). Changes in premenstrual symptoms and irrational thinking following cognitive–behavioral coping skills training. *Journal of Consulting and Clinical Psychology, 62,* 1026–1032.

Lambert, M. J., & Bergin, A. E. (1994). The effectiveness of psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 143–189). New York: Wiley.

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.

*Markowitz, J. C., Kocsis, J. H., Fishman, B., Spielman, L. A., Jacobsberg, L. B., Frances, A. J., et al. (1998). Treatment of depressive symptoms in human immunodeficiency virus-positive patients. *Archives of General Psychiatry, 55,* 452–457.

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

*O'Malley, S. S., Jaffe, A. J., Chang, G., Schottenfeld, R. S., Meyer, R. E., & Rounsaville, B. (1992). Naltrexone and coping skills therapy for alcohol dependence. *Archives of General Psychiatry, 49,* 881–887.

*Pinto, A., La Pia, S., Mennella, R., Giorgio, D., & DeSimone, L. (1999). Cognitive–behavioral therapy and clozapine for clients with treatment-refractory schizophrenia. *Psychiatric Services, 50,* 901–904.

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.

*Sensky, T., Turkington, D., Kingdon, D., Scott, J. L., Scott, J., Siddle, R., et al. (2000). A randomized controlled trial of cognitive–behavioral therapy for persistent symptoms in schizophrenia resistant medication. *Archives of General Psychiatry, 57,* 165–172.

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

*Shear, M. K., Pilkonis, P. A., Cloitre, M., & Leon, A. C. (1994). Cognitive–behavioral treatment compared with nonprescriptive treatment of panic disorder. *Archives of General Psychiatry, 51,* 395–401.

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

*Stanley, M. A., Beck, J. G., & Glassco, J. D. (1996). Treatment of generalized anxiety in older adults: A preliminary comparison of cognitive–behavioral and supportive approaches. *Behavior Therapy, 27,* 565–581.

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.

*Tarrier, N., Yusupoff, L., Kinney, C., McCarthy, E., Gledhill, A., Haddock, G., & Morris, J. (1998). Randomised controlled trial of intensive cognitive behaviour therapy for patients with chronic schizophrenia. *British Medical Journal, 317,* 303–307.

*Thase, M. E., Friedman, E. S., Berman, S. R., Fasiczka, A. L., Lis, J. A., Howland, R. H., & Simons, A. D. (2000). Is cognitive behavior therapy just a 'nonspecific' intervention for depression? A retrospective comparison of consecutive cohorts treated with cognitive behavior therapy or supportive counseling and pill placebo. *Journal of Affective Disorders, 57,* 63–71.

*Treasure, J., Todd, G., Brolly, M., Tiller, J., Nehmed, A., & Denman, F.

(1995). A pilot study of a randomised trial of cognitive analytical therapy vs. educational behavioral therapy for adult anorexia nervosa. *Behavior Research and Therapy, 33,* 363–367.

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,* 1–13.

*Wilson, G. T., Loeb, K. L., Walsh, B. T., Labouvie, E., Petkova, E., Liu, X., & Waternaux, C. (1999). Psychological versus pharmacological treatments of bulimia nervosa: Predictors and processes of change. *Journal of Consulting and Clinical Psychology, 67,* 451–459.

Received March 27, 2003
Revision received June 13, 2003
Accepted June 18, 2003 ■