
Hyun-nie Ahn and Bruce E. Wampold
University of Wisconsin—Madison

Component studies, which involve comparisons between a treatment package and the treatment package without a theoretically important component or the treatment package with an added component, use experimental designs to test whether the component is necessary to produce therapeutic benefit. A meta-analysis was conducted on 27 component studies culled from the literature. It was found that the effect size for the difference between a package with and without the critical components was not significantly different from zero, indicating that theoretically purported important components are not responsible for therapeutic benefits. Moreover, the effect sizes were homogeneous, which suggests that there were no important variables moderating effect sizes. The results cast doubt on the specificity of psychological treatments.

It was established in the 1980s that counseling and psychotherapy are remarkably efficacious (Lambert & Bergin, 1994; Wampold, 2000); now on center stage is the controversy about whether the beneficial effects of counseling and psychotherapy are due to the specific ingredients of the treatments or to the factors common in all therapies (Wampold, 2000). On one side are the advocates of empirically supported treatments who claim that treatments are analogues of medical treatments in that efficacy is attributed to their respective specific ingredients, which are usually presented in treatment manuals (see, e.g., Chambless & Hollon, 1998; Chambless et al., 1996; Crits-Christoph, 1997; DeRubeis & Crits-Christoph, 1998; DeRubeis et al., 1990; DeRubeis & Feeley, 1990; Task Force on Promotion and Dissemination of Psychological Procedures, 1995; Waltz, Addis, Koerner, & Jacobson, 1993; Wilson, 1996). Specificity (i.e., attributing outcome to specific ingredients) is one of the hallmarks of the medical model. On the other side are the advocates of models that stipulate that the common factors, such as the healing context, the working alliance, and belief in the rationale for treatment and in the treatment itself, are the important therapeutic aspects of counseling and psychotherapy (see, e.g., Frank & Frank, 1991; Garfield, 1992; Luborsky, Singer, & Luborsky, 1975; Parloff, 1986; Rosenzeig, 1936; Strupp, 1986; Wampold, 1997, 2000, 2001; Wampold et al., 1997). From a scientific perspective, the specific ingredient versus common factor polemic should be settled empirically rather than rhetorically.

Demonstrating that the specific ingredients of a treatment are responsible for the benefits of counseling and psychotherapy is complex (see Wampold, 2001, for a discussion of research strategies for establishing specificity). There are many research strategies that can be used to demonstrate the specificity of psychological treatments. Of such designs, component studies come closest to the "gold standard" of experimental designs and, as such, should show evidence for specificity, should specificity exist. Component studies attempt to isolate the effects of ingredients by comparing treatments with and without those ingredients. Component studies contain two similar designs, dismantling designs and additive designs.

One crucial feature of the [dismantling] design is that more factors are poorly outcomes vis-à-vis the complete treatment, evidence accrues for the specificity of those ingredients. Borkovec (1990) described the advantages of the dismantling study:

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. (pp. 56–57)
This study illustrates the logic of the component design. As well, combined with coping skills related to automatic thoughts, or the three components: behavioral activation, coping strategies for efficacy. Jacobson et al. separated cognitive-behavioral therapy into (1996) to determine what components of cognitive-behavioral dysfunctional schemas. Generally, the results showed equivalence, coping skills, and identification and modification of core depressogenic cognition. Participants were randomly assigned to a behavioral activation group, a treatment involving behavioral activation combined with coping skills related to automatic thoughts, or the complete cognitive treatment, which included behavioral activation, coping skills, and identification and modification of core dysfunctional schemas. Generally, the results showed equivalence in outcomes across the groups at termination and at follow-up. This study illustrates the logic of the component design. As well, the results failed to produce evidence of the specificity of ingredients of cognitive–behavioral therapy.

If specific ingredients are indeed responsible for the benefits of counseling and psychotherapy, then component studies should consistently demonstrate an effect when a treatment condition is compared with a condition not involving a theoretically stipulated component. Bearing in mind that a few component studies could demonstrate such differences by chance (i.e., Type I errors), it is important to determine whether the corpus of component studies produces specificity effects. Meta-analysis has been shown to be a powerful method to review literature and bring clarity to disputes in education, medicine, psychology, and public policy (Hunt, 1997; Mann, 1994). The purpose of this study was to meta-analytically examine component studies to determine the degree to which these studies produce evidence that supports the specificity of psychological treatments.

Method

Because this meta-analysis involved a methodological feature (viz., component studies), determining a keyword for an electronic literature search was not possible. Therefore, a comprehensive search of journals that publish outcome research was undertaken. Wampold et al. (1997) reviewed the research included in Shapiro and Shapiro’s (1982) meta-analysis of comparative studies and found that the preponderance of such studies were published in four journals: Behaviour Research and Therapy, Behavior Therapy, Journal of Consulting and Clinical Psychology, and Journal of Counseling Psychology. Sues, Shapiro, and Elliott (1986) noted that detecting the relative efficacy of treatments depended on sophisticated research methods and that more recent studies, involving improved methods, would be more likely to reveal differences between treatments, should they be present. Accordingly, we searched for component studies published in the most recent decade (i.e., 1990 to 1999) in the four identified journals. This strategy eliminated dissertations, presentations, and other unpublished studies. However, given that studies with statistically significant results are more likely to be published (Atkinson, Furlong, & Wampold, 1982), omitting unpublished studies would tend to overestimate the effect of specific ingredients; consequently, the present analysis yields a liberal test of specificity.

In identifying the studies for this meta-analysis, Hyun-nie Ahn examined every study published in the four journals just identified from 1990 to 1999. To be included in this meta-analysis, a study had to (a) involve a psychological treatment intended to be therapeutic for a particular disorder, problem, or complaint and (b) contain the necessary statistics to conduct the meta-analysis. To determine that a treatment was intended to be therapeutic, we used the criteria developed by Wampold et al. (1997); specifically, a treatment had to involve a therapist who had at least a master’s degree and who met face to face with the client and developed a relationship with the client. Moreover, the treatment had to contain at least two of the following four elements: (a) The treatment was based on an established treatment that was cited, (b) a description of the treatment was contained in the article, (c) a manual was used to guide administration of the treatment, and (d) active ingredients of the treatment were identified and cited. Finally, the study’s research design had to involve a comparison of one group with another group, and one of the following two conditions had to be satisfied: (a) One, two, or three ingredients of the treatment were removed, leaving a treatment that would be considered logically viable (i.e., coherent and credible), or (b) one, two, or three ingredients that were compatible with the whole treatment and were theoretically or empirically hypothesized to be active were added to the treatment, providing a “super treatment.” A study was excluded when treatment A was compared with...
treatment B, where B was a subset of A but both A and B were established treatments in their own rights.

Initially, all studies were gathered that compared one treatment group with another group that had components added or removed, although the study may not have met the inclusion and exclusion criteria. Two raters (both doctoral students in counseling psychology) were then asked to determine the suitability of each study for this meta-analysis using a rating sheet listing the inclusion and exclusion criteria. A study was retained if both raters agreed on its inclusion in the study. When the raters disagreed on a study, Bruce E. Wampold rated the study, and the study was included if he determined that it met the criteria. The resulting meta-analytic sample included 27 treatment comparisons derived from 20 studies (see Table 1).

**Analytic Strategy**

For each study i, an estimate of the effect size $d_i$ for study i that reflected the effect of a given component or components, as well as an estimate of the variance of this estimate—that is, $\hat{\sigma}^2(d_i)$—was calculated in the following way. First, for each dependent variable, a sample effect size was obtained by calculating the difference in the means of the two conditions and standardizing by dividing by the pooled standard deviation: (more-component-group $M - \text{fewer-component-group } M)/SD$. This value was adjusted to yield an unbiased estimate of the population effect size; as well, the standard error of estimate was calculated (Hedges & Olkin, 1985). To determine a single estimate of the effect size for each study, we combined the effect sizes for each dependent variable under the assumption that the correlation among the dependent variables was .50, a reasonable value for this correlation in psychotherapy studies (see Hedges & Olkin, 1985, pp. 212–213, for the method and Wampold et al., 1997, for a justification and application in the psychotherapy context). This procedure yielded, for study $i$, the desired estimates $d_i$ and $\hat{\sigma}^2(d_i)$; it also provided a more precise estimate of $d_i$ (i.e., reduced the standard error of estimate) than would the estimate for any single dependent variable (Wampold et al., 1997).

To aggregate the effect sizes over the 27 comparisons, we weighted each $d_i$ by the inverse of the variance, in the standard fashion, to yield the aggregated effect size estimate $d$ (Hedges & Olkin, 1985). As well, the

<table>
<thead>
<tr>
<th>Study</th>
<th>Disorder</th>
<th>More components group</th>
<th>Fewer components group</th>
<th>Component(s) tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appelbaum et al. (1990)</td>
<td>Tension headache</td>
<td>CT + PMR</td>
<td>PMR</td>
<td>Cognitive component</td>
</tr>
<tr>
<td>Barlow et al. (1992)</td>
<td>Generalized anxiety disorder</td>
<td>CT + PMR</td>
<td>CT</td>
<td>Relaxation skills</td>
</tr>
<tr>
<td>Baucom et al. (1990)</td>
<td>Marital discord</td>
<td>CR + BMT</td>
<td>PMR</td>
<td>CR</td>
</tr>
<tr>
<td>Blanchard et al. (1990)</td>
<td>Tension headache</td>
<td>EET + BMT</td>
<td>BMT</td>
<td>EET</td>
</tr>
<tr>
<td>Borkovec &amp; Costello (1993)</td>
<td>Generalized anxiety disorder</td>
<td>CT + PMR</td>
<td>PMR</td>
<td>Cognitive component</td>
</tr>
<tr>
<td>Blanchard et al. (1990)</td>
<td>Generalized anxiety disorder</td>
<td>CT + PMR</td>
<td>PMR</td>
<td>Cognitive component</td>
</tr>
<tr>
<td>Dadds &amp; McHugh (1992)</td>
<td>Child conduct problem</td>
<td>CMT + EET</td>
<td>CMT</td>
<td>Social support</td>
</tr>
<tr>
<td>Defenbacher &amp; Stark (1992)</td>
<td>General anger</td>
<td>CRCS</td>
<td>RCS</td>
<td>Cognitive component</td>
</tr>
<tr>
<td>Feske &amp; Goldstein (1997)</td>
<td>Panic disorder</td>
<td>EMDR</td>
<td>EFER</td>
<td>Eye movement</td>
</tr>
<tr>
<td>Halford et al. (1993)</td>
<td>Marital discord</td>
<td>Enhanced BMT</td>
<td>BMT</td>
<td>CR + generalized training + affective exploration</td>
</tr>
<tr>
<td>Hope et al. (1995)</td>
<td>Social phobia</td>
<td>CBT</td>
<td>Exposure only</td>
<td>Cognitive component</td>
</tr>
<tr>
<td>Jacobson et al. (1996)</td>
<td>Depression</td>
<td>BA + AT</td>
<td>AT</td>
<td>BA</td>
</tr>
<tr>
<td>Nicholas et al. (1991)</td>
<td>Chronic low back pain</td>
<td>CT + PMR</td>
<td>CT</td>
<td>Relaxation skills</td>
</tr>
<tr>
<td>Öst et al. (1991)</td>
<td>Blood phobia</td>
<td>Applied tension package (BT)</td>
<td>BT</td>
<td>Behavioral component</td>
</tr>
<tr>
<td>Porzelius et al. (1995)</td>
<td>Eating disorder</td>
<td>OBET</td>
<td>Applied tension package (BT)</td>
<td>Exposure in vivo</td>
</tr>
<tr>
<td>Propst et al. (1992)</td>
<td>Depression</td>
<td>CBT-Religious</td>
<td>CBT</td>
<td>Advanced CBT with a focus on coping skills and cognitive interventions</td>
</tr>
<tr>
<td>Radojevic et al. (1992)</td>
<td>Rheumatoid arthritis</td>
<td>BT + social support</td>
<td>BT</td>
<td>Religious content modified to fit CBT</td>
</tr>
<tr>
<td>Rosen et al. (1990)</td>
<td>Body image</td>
<td>CBT + size perception training</td>
<td>CBT</td>
<td>Family support</td>
</tr>
<tr>
<td>Thackwray et al. (1993)</td>
<td>Bulimia nervosa</td>
<td>CBT</td>
<td>GDVM</td>
<td>Cognitive component</td>
</tr>
<tr>
<td>Webster-Stratton (1994)</td>
<td>Parenting effectiveness</td>
<td>GDVM + ADVANCE</td>
<td>BT</td>
<td>Cognitive social learning + group discussion</td>
</tr>
<tr>
<td>Williams &amp; Falbo (1996)</td>
<td>Panic attack with agoraphobia</td>
<td>CBT</td>
<td>BT</td>
<td>Cognitive component</td>
</tr>
</tbody>
</table>

Note. CT = cognitive therapy; PMR = progressive muscle relaxation; CR = cognitive restructuring; BMT = behavioral marital therapy; EET = emotional expressiveness training; CBT = cognitive–behavioral therapy; AR = applied relaxation; CMT = child management training; CRCS = cognitive and relaxation coping skills; RCS = relaxation coping skills; EMDR = eye movement desensitization and reprocessing; EFER = eye fixation exposure and reprocessing; BA = behavioral activation; AT = automatic thoughts; BT = behavioral therapy; OBET = obese binge eating treatment; GDVM = videotaped parent skills training program; ADVANCE = cognitive training social learning program.
standard error of this estimate \((d_+\)\), which is used to calculate the confidence interval of \(d_+\) and to test the null hypothesis that the population effect size is zero, was calculated according to the methods developed by Hedges and Olkin. Finally, a homogeneity test was conducted to determine whether the 20 effect sizes were drawn from the same population.

Results

Using the aggregation strategy just described, we obtained the following estimates: \(d_+ = -0.20\) and \(\sigma^2(d_+) = 0.176\). The negative value for \(d_+\) indicates that the treatment conditions with fewer components outperformed the treatment conditions with more components, a result in the opposite direction from that anticipated. In any event, an effect size of magnitude 0.20 is considered small (Cohen, 1988).

The 95% confidence interval for the population effect size, given a normal effect size distribution, was as follows: lower bound, \(d_+ - 1.96 \sigma(d_+) = -0.541\), and upper bound, \(d_+ + 1.96 \sigma(d_+) = 0.149\). Because this confidence interval contained zero, the null hypothesis that the population effect size is zero was not rejected.

To determine whether the effect sizes for the 20 comparisons were drawn from a single population, we conducted a test of homogeneity using the methods described by Hedges and Olkin (1985). The \(Q\) statistic is a goodness-of-fit statistic, as follows:

\[
Q = \sum_{i=1}^{k} \frac{(d_i - d_+)^2}{\sigma^2(d_i)},
\]

where \(k\) is the number of studies aggregated. The \(Q\) statistic has approximately a chi-square distribution with \(k - 1\) degrees of freedom. If \(Q\) is sufficiently large, the homogeneity hypothesis is rejected. In the present case, \(Q = 33.34\), which, when compared with a chi-square distribution with 26 degrees of freedom, was insufficiently large to reject the null; therefore, it was concluded that the effect sizes were homogeneous. Thus, it appears that there were no variables that would moderate the overall effect size, which was not different from zero. However, this conclusion must be tempered by the fact that the power of the homogeneity test can be low when various assumptions are violated and the sample sizes of the studies are small in comparison with the number of studies (see Harwell, 1997).

Discussion

The present meta-analysis of component studies produced no evidence that the specific ingredients of psychological treatments are responsible for the beneficial outcomes of counseling and psychotherapy. For example, the aggregate effect size for comparisons was not significantly different from zero. Moreover, the effect sizes from the 27 comparisons were homogeneous, ruling out rival hypotheses that a missing variable would moderate the relationship between components and outcome.

It should be recognized that the studies reviewed in this meta-analysis examined treatments that have been found to be efficacious. Moreover, the component removed or added was hypothesized by the researchers to be efficacious according to the theoretical tenets of the respective treatments. For example, in the component study described in the introduction, Jacobson et al. (1996) clearly described the theoretical basis of the study:

Beck and his associates are quite specific about the hypothesized active ingredients of CT [cognitive–behavioral treatment], stating throughout their treatment manual (Beck et al., 1979) that interventions aimed at cognitive structures or core schema are the active change mechanisms [for treating depression]. Despite this conceptual clarity, the treatment is so multifaceted that a number of alternative accounts for its efficacy are possible. We label two primary competing hypotheses the “activation hypothesis” and the “coping skills” hypothesis... If an entire treatment based on activation interventions proved to be as effective as CT, the cognitive model of change in CT (stipulating the necessary interventions for the efficacy of CT) would be called into question. (pp. 295–296)

In the Jacobson et al. (1996) study, the authors were examining the most validated psychotherapeutic treatment in existence, namely cognitive–behavioral treatment for depression, and testing whether the cognitive ingredients were indeed necessary to produce benefits.

A criticism could be raised that included in the corpus of studies examined were some ingredients that are important and others that are not and that aggregating across diverse studies yields spurious conclusions. This is a familiar criticism of meta-analysis. First, the homogeneity finding suggests that there are not two classes of comparisons, those with efficacious specific ingredients and those without. Second, an occasional study demonstrating that a component was related to the outcome must be considered, in light of the present results, a Type I error. The argument that a given specific ingredient is efficacious would need to be supported by replications, a situation not evident in the studies reviewed. Third, it is important to note that Jacobson et al.’s dismantling of the empirically supported cognitive–behavioral treatment of depression, probably the most established psychological treatment in existence, failed to demonstrate that the components of the treatment were responsible for the benefits.

The evidence produced by this meta-analysis casts suspicion on the specificity of psychological treatments. Although some of the treatments contained in the studies reviewed were designed for disorders that are not prevalent (e.g., blood phobia), all of the treatments contained discrete components that lend themselves to detecting the efficacy of specific ingredients. That is, if the specific ingredients of treatments are responsible for the benefits of psychotherapy, then the expected effects should appear in the studies reviewed. As well, it would not be expected that specific ingredients of treatments with less well-defined components would be responsible for the benefits of such treatments.

Other research evidence tends not to support the benefits of specific ingredients of psychological treatments. If specific ingredients were remedial for a problem, then it would be expected that some treatments (viz., those containing potent specific ingredients) would be superior to other treatments. However, the outcome research conclusively has shown that all treatments produce approximately equal benefits generally (Wampold, 2000; 2001; Wampold et al., 1997) as well as in particular areas, such as depression (e.g., Elkin et al., 1989; Robinson, Berman, & Neimeyer, 1990; Wampold, Minami, Baskin, & Tierney, in press) and anxiety (see Wampold, 2001). Attempts to demonstrate specificity by examining mediating effects have failed to show that specific
treatments work through the theoretically hypothesized mechanisms (Wampold, 2001). For example, in the National Institute of Mental Health Treatment of Depression Collaborative Research Program, cognitive–behavioral treatment and interpersonal treatments did not operate uniquely through the intended respective cognitive and interpersonal mechanisms, as hypothesized (Imber et al., 1990). Finally, specificity predicts that certain treatments will be particularly effective with clients with certain deficits, for example, cognitive treatments for clients with irrational thoughts and interpersonal treatments for clients with maladaptive social relations. However, theoretically predicted interactions between treatments and client characteristics of this type have never been demonstrated (for laudable attempts, see McKnight, Nelson-Gray & Barnhill, 1992; Project MATCH Research Group, 1997; Simons, Garfield, & Murphy, 1984).

The results of the present meta-analytic study are not an anomaly in an otherwise uniform field of research results supporting specificity; rather, the preponderance of the research evidence is not supportive of the benefits of specific ingredients. This suggests that the benefits of treatments are probably due to the pathways common to all bona fide psychological treatments, such as the healing context, the belief in the rationale for and the efficacy of therapy by the client and by the therapist, the therapeutic alliance, therapeutic procedures consistent with the client’s understanding of his or her problems, the development of increased self-efficacy to solve one’s problems, and remoralization (Frank & Frank, 1991; Garfield, 1992; Wampold, 2001). The research evidence supports the notion that the benefits of counseling and psychotherapy are derived from the common factors. For example, it has been shown that the therapeutic alliance, measured at an early stage, accounts for a significant portion of the variability in treatment outcomes (Horvath & Symonds, 1991; Martin, Garske, & Davis, 2000). Moreover, the variance due to therapists within treatments is greater than the variance between treatments, lending primacy to the person of the therapist rather than to the particular treatment (Crits-Christoph et al., 1991; Wampold & Serlin, 2000). Indeed, the common factors account for about 9 times more variability in outcomes than do the specific ingredients (Wampold, 2001).

Rejecting the specificity of counseling and psychotherapy has implications for training, practice, and research. Training models should focus on the common factors as the bedrock of skills necessary to become an effective practitioner. The importance of interviewing skills, establishment of a therapeutic relationship, and the core facilitative conditions in the training of counselors and psychologists is supported by the empirical evidence. Omitting these vital components and training students to conduct solely various empirically supported treatments is contraindicated. Nevertheless, counselors and therapists need to learn techniques, a position well stated by common factor advocate Jerome Frank:

"My position is not that technique is irrelevant to outcome. Rather, I maintain that, as developed in the text, the success of all techniques depends on the patient’s sense of alliance with an actual or symbolic healer. This position implies that ideally therapists should select for each patient the therapy that accords, or can be brought to accord, with the patient’s personal characteristics and view of the problem. Also implied is that therapists should seek to learn as many approaches as they find congenial and convincing. Creating a good therapeutic match may involve both educating the patient about the therapist’s conceptual scheme and, if necessary, modifying the scheme to take into account the concepts the patient brings to therapy. (Frank & Frank, 1991, p. xv)

The use of treatment manuals in practice is not supported by the research evidence. Although standardization of treatment appears scientific and may be required for experimental control in the research context, there is no evidence that adherence to a treatment protocol results in superior outcomes; in fact, slavish adherence to a manual can cause ruptures in the alliance and, consequently, poorer outcomes (Wampold, 2001). As well, use of manuals restricts adaptation of treatments to the attitudes, values, and culture of the client, a necessary aspect of multicultural counseling.

A common factor perspective places emphasis on the skill of the therapist. There is compelling evidence that a large proportion of variability in outcomes is due to therapists, even when therapists are “experts” in a particular approach and are supervised and monitored (Wampold, 2001, chap. 8). Thus, emphasis should be placed on the therapist or counselor rather than on the particular therapy. Consequently, those who control access to therapy (e.g., health maintenance organizations) should refer clients to counselors who have demonstrated efficacy rather than mandate particular services. Indeed, it would be in the best interest of agencies to have therapists of various orientations so that clients could receive the type of therapy that best accords with their worldview.

Combined with the evidence that all bona fide treatments are equally efficacious (see Wampold, 2001, chap. 4), the results of this meta-analysis suggest that comparative outcome studies will yield nonsignificant differences and therefore are costly experiments in futility. It is safe to say that hundreds of millions of dollars have been spent on outcome research that has shown that bona fide psychological treatments are efficacious but that all such treatments produce about the same benefits. Continued outcome research will only support that general pattern of results and yield little informative evidence about counseling and psychotherapy. Rather, the focus of counseling research should be on the process of counseling and on the common factors that have historically interested humanistic and dynamic researchers and clinicians.

References

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


Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that "everyone has won and all must have prizes?" *Archives of General Psychiatry*, 32, 995–1008.


WHERE ARE THE SPECIFIC INGREDIENTS?


Received July 14, 2000
Revision received September 26, 2000
Accepted October 24, 2000