Therapist Adherence/Competence and Treatment Outcome: A Meta-Analytic Review

Christian A. Webb,
Department of Psychology, University of Pennsylvania, Philadelphia.

Robert J. DeRubeis, and
Department of Psychology, University of Pennsylvania, Philadelphia.

Jacques P. Barber
Department of Psychiatry, University of Pennsylvania Medical School, and Department of Psychiatry, Veterans Affairs Medical Center, Philadelphia.

Abstract

Objective—The authors conducted a meta-analytic review of adherence–outcome and competence– outcome findings, and examined plausible moderators of these relations.

Method—A computerized search of the PsycINFO database was conducted. In addition, the reference sections of all obtained studies were examined for any additional relevant articles or review chapters. The literature search identified 36 studies that met the inclusion criteria.

Results—R-type effect size estimates were derived from 32 adherence–outcome and 17 competence–outcome findings. Neither the mean weighted adherence– outcome ($r = .02$) nor competence–outcome ($r = .07$) effect size estimates were found to be significantly different from zero. Significant heterogeneity was observed across both the adherence–outcome and competence–outcome effect size estimates, suggesting that the individual studies were not all drawn from the same population. Moderator analyses revealed that larger competence–outcome effect size estimates were associated with studies that either targeted depression or did not control for the influence of the therapeutic alliance.

Conclusions—One explanation for these results is that, among the treatment modalities represented in this review, therapist adherence and competence play little role in determining symptom change. However, given the significant heterogeneity observed across findings, mean effect sizes must be interpreted with caution. Factors that may account for the nonsignificant adherence– outcome and competence–outcome findings reported within many of the studies reviewed are addressed. Finally, the implication of these results and directions for future process research are discussed.

Keywords

meta-analysis; therapist; adherence; competence

Numerous forms of psychotherapy have been developed. These treatments vary in the extent to which they have garnered empirical support in their favor (Chambless & Ollendick, 2001; Klonsky, 2009). Although evidence has accrued in support of the efficacy of certain interventions, the mechanisms through which these therapies exert their beneficial effects are generally not well understood (Kazdin, 2006).

Researchers have hypothesized a number of different "active ingredients" that may be responsible for the therapeutic improve ment that many patients receiving these treatments enjoy. Traditionally, plausible active ingredients have been classified into two broad categories, *common factors* and *specific factors* (Castonguay, 1993). The term *common factors* refers to those elements that are shared across most, if not all, therapeutic modalities (e.g., a convincing rationale, expectations of improvement, a therapeutic alliance). The alliance has received substantial attention within the psychotherapy literature and has been examined in relation to outcome across a variety of treatment modalities and mental health problems. Reviews of this literature indicate that a stronger alliance is associated with better treatment outcomes. For example, in their meta-analysis, Martin, Garske, and Davis (2000) reported a mean alliance–outcome correlation of .22.

Although *specific factors* have generally been described in contrast to *common factors*, the former term can be misleading in that it often refers to techniques or methods that are not specific to a particular treatment but rather are employed in many treatment modalities. For example, some of the cognitive techniques used in cognitive therapy are also employed within other forms of psychotherapy. Rather than referring to elements that are unique to a particular orientation, the term *specific factors* is generally meant to refer to the core, *theory-specified* techniques or methods that are prescribed for a given treatment modality (Castonguay & Holt-forth, 2005). To return to the cognitive therapy example, cognitive techniques intended to help patients identify and challenge mal-adaptive thoughts are central components of the treatment and, according to cognitive therapy theory, play a key role in contributing to symptom improvement (DeRubeis, Webb, Tang, & Beck, 2009). Although such cognitive strategies may be employed within other treatment modalities, they are generally not specified as the core therapeutic ingredients in the theories of change associated with these treatments.

Among manualized therapies, there has been an interest both in the degree to which therapists are delivering the theory-specified techniques or methods of the intervention (therapist *adherence*), and the skill with which these techniques or methods are implemented (therapist *competence*; Barber et al., 2006; Sharpless & Barber, 2009). There are several reasons for the interest in therapist adherence and competence. First, within treatment outcome studies, measures of adherence and competence have been employed in an effort to monitor *treatment integrity*. Specifically, there is an interest in evaluating the degree to which study therapists are adhering to the treatment under investigation, examining whether they are delivering its techniques or methods in a competent manner, and, if multiple

treatment modalities are being evaluated, ensuring that these treatments can be differentiated from one another (Perepletchikova & Kazdin, 2005). For example, in their randomized placebo-controlled trial, Dimidjian et al. (2006) used the Cognitive Therapy Scale (CTS; Young & Beck, 1980) to assess the competence with which their cognitive therapy condition was implemented. Dimidjian et al. employed the frequently used criterion of a total CTS score of 40 or greater as their threshold (or redline criterion) of acceptable competence in cognitive therapy. In the Treatment of Depression Collaborative Research Program (TDCRP; Elkin et al., 1989), Hill, O'Grady, and Elkin (1992) assessed therapist adherence to cognitive–behavioral therapy, interpersonal therapy, and clinical management and demonstrated that these three modalities could be differentiated reliably using the Collaborative Study Psychotherapy Rating Scale (CSPRS; Hollon et al., 1988).

In addition, within psychotherapy process research, there is an interest in examining the statistical relation between variability in ratings of adherence and competence and scores on outcome measures. Such studies have been conducted in an effort to identify the active ingredients within particular treatments, with the ultimate goal of improving treatment efficacy. That is, with a better understanding of which elements contribute to outcome, researchers can modify treatments, providing the optimal dose of active ingredients and minimizing inert or iatrogenic elements (Kazdin, 2006).

Ideally, studies in which process–outcome relations are examined would take the form of well-controlled experiments in which one therapy process variable is manipulated (e.g., therapist use of concrete cognitive therapy techniques), while all others are held constant and patients are randomly assigned to conditions (see Høglend et al., 2008, for a recent example of an experimental study of transference interpretations in dynamic psychotherapy). However, researchers examining the relation between therapist adherence or competence and outcome have generally employed observational, rather than experimental, methods. In most cases, trained raters have coded one or more videotaped or audiotaped therapy sessions using measures of therapist adherence or competence, and scores on these measures have been correlated with posttreatment scores on an outcome measure (generally with statistical controls for pretreatment scores on the given outcome measure).

In order for observational studies to provide strong support for a causal claim, three criteria should be satisfied: (a) covariation between variables, (b) nonspuriousness, and (c) the temporal precedence of cause before effect (Judd & Kenny, 1981). Although a significant process–outcome correlation may be reported, such an association may be spurious. If a significant correlation is obtained between therapist adherence to concrete cognitive therapy techniques and change in depressive symptoms (e.g., Webb et al., 2009), it is possible that one or more unmeasured third-variable confounds account for this association. For example, certain patient characteristics may lead therapists to deliver higher levels of concrete cognitive therapy techniques, while also contributing to improvement in depressive symptoms. Although third-variable confounds are an inherent limitation of observational data, investigators in some studies of adherence–outcome (e.g., DeRubeis & Feeley, 1990; Feeley, DeRubeis, & Gelfand, 1999) and competence–outcome (e.g., Barber, Crits-Christoph, & Luborsky, 1996; Svartberg & Stiles, 1994; Trepka, Rees, Shapiro, Hardy, & Barkham, 2004) relations have attempted to control for at least one plausible third variable,

most often the alliance. Furthermore, adherence or competence may interact with other variables to contribute to symptom improvement. For example, Barber et al. (2006) found that there was a quadratic relation between adherence to individual drug counseling and outcome and that this quadratic relation was moderated by alliance level. This finding also exemplifies the possibility that the relation between adherence or competence and outcome may, at least in some contexts, be nonlinear, underscoring the importance of examining curvilinear relations.

Testing for third-variable confounds has been more common than has the attempt to rule out the possibility that the statistical association between process and outcome derives from the effect of outcome on process. The vast majority of studies of alliance–outcome relations included in the Martin et al. (2000) meta-analysis, for example, employed methods that cannot be used to identify the direction of any causal effect between the alliance and symptom change. Studies of adherence or competence have also often failed to control for temporal confounds.

There have been fewer published studies of the relation between adherence or competence and outcome, relative to the numerous studies of alliance–outcome relations. This is likely due, at least in part, to the common factors *zeitgeist*, fueled by proponents of the "Dodo bird verdict" (e.g., Wampold, 2001). Another factor that may help account for the relative paucity of adherence–outcome and competence–outcome studies is that the methodologies most often employed to assess therapist adherence and competence are time and labor intensive. Rather than relying on reports from therapists or patients (as is often the case for assessments of the alliance; Elvins & Green, 2008; Martin et al. 2000), adherence and competence studies usually involve raters trained to code these variables from videotapes or audiotapes of therapy sessions.

To our knowledge, no meta-analysis of either adherence–outcome or competence–outcome relations has been published. To fill this void, for our meta-analytic review we aggregated studies in which therapist adherence or competence was examined in relation to outcome. As we discuss later in more detail, in addition to including studies in which the relation between ratings of entire adherence or competence scales and outcome were examined, we also included studies that assessed only particular components of adherence or competence that were believed to be especially important to the treatment modality in question on the basis of either prior empirical evidence or theory (e.g., Intrapersonal Consequences subscale of the Coding System of Therapist Feedback [CSTF] for Castonguay, Goldfried, Wiser, Raue, & Hayes, 1996; various measures assessing therapist interpretations in dynamic therapy for Crits-Christoph, Cooper, & Luborsky, 1988; Marziali, 1984; Piper, Azim, Joyce, & McCallum, 1991; Piper, Debbane, Bienvenu, de Carufel, & Garant, 1986; the Errors in Technique subscale of the Vanderbilt Negative Indicators Scale [VNIS] for Sachs, 1983; Cognitive Therapy–Concrete subscale of the CSPRS for Webb et al., 2009).

It may be that variability in adherence or competence is significantly related to variability in outcome in some contexts but not others (e.g., only within some treatment modalities). To the extent that this is the case, pooling adherence–outcome and competence–outcome effect size estimates across studies that vary on a number of different characteristics may be

unwise, yielding mean effect sizes that are relatively uninformative (Lipsey & Wilson, 2001). However, as a starting point, effect sizes were initially pooled across studies, and mean weighted adherence–outcome and competence–outcome effect sizes were estimated. This approach has also been employed in quantitative reviews of alliance–outcome relations (Horvath & Symonds, 1991; Martin et al., 2000), allowing for informative comparisons of the present findings with findings from meta-analyses of alliance–outcome studies.

In addition, we conducted homogeneity analyses to test whether the constituent effect size estimates were drawn from the same population and to quantify the degree of heterogeneity across effect size estimates. Moreover, as discussed later in more detail, we coded the studies on a variety of dimensions that represent plausible moderators of adherence–outcome and competence– outcome relations.

Method

Inclusion Criteria

The current study's inclusion criteria were based, in part, on the criteria employed in the Martin et al. (2000) meta-analysis of alliance–outcome relations. Specifically, in order to have been included in the current meta-analysis, a study must have possessed the following characteristics: (a) it was an investigation of individual in-person psychotherapy, as opposed to group, family, or couples treatment; (b) it included a quantifiable measure of both adherence or competence and outcome (assessed subsequently and no later than 6 months after treatment termination); (c) its assessments of adherence or competence had to have been based on videotaped, audiotaped, or transcribed therapy sessions rated by experts or trained raters, rather than by therapists or patients; (d) it had to include a clinical rather than analogue population; (e) it must have comprised at least five patients in (each of) the treatment group(s); and (f) the publication must have been in English.

Literature Search Procedures

Several literature search techniques were used to identify applicable studies for inclusion. First, a computerized search of the PsycINFO database was conducted. Studies with any possible combination of the keywords *therapist*, *therapy*, *psychotherapist*, *psychotherapy*, and *adherence*, *adhering*, *adherent*, *adhere*, *competence*, *competent*, *competency*, *competencies*, *integrity* were examined. The resulting list of 5,868 articles was inspected for relevant studies meeting the aforementioned inclusion criteria. The final PsycINFO database search was completed April 15, 2009. In addition, we searched the obtained studies and review chapters, as well as the reference sections of these sources for any additional relevant articles. In order to obtain all relevant studies and limit publication bias, we also included unpublished dissertations.

The literature search identified 36 studies that met the inclusion criteria (designated in the reference list by an asterisk). Given that several of these studies examined both adherence– outcome and competence–outcome relations, or explored process–outcome relations within multiple treatment modalities, a total of 49 distinct effect size estimates (32 adherence– outcome and 17 competence– outcome values) were extracted from these studies and

included in the current meta-analysis.[1] The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA, formerly Quality of Reporting of Meta-Analyses, or QUOROM) flow diagram is included in the online supplemental materials (Moher, Liberati, Tetzlaff, Altman, & the PRISMA Group, 2009).

Meta-Analytic Procedures

Among the obtained studies, the most commonly reported effect size estimate of the relation between therapist adherence (or competence) and outcome was a *r*-type effect size estimate, most often with a statistical control for pretreatment scores on the given outcome measure. Thus, the *r*-type effect size statistic was selected as the effect size metric for the current meta-analysis. For studies that did not report adherence–outcome or competence–outcome *r*-type effect size estimates, the authors were contacted directly in an effort to obtain this effect size estimate format. If the authors were unable to provide the data, or did not respond, we computed *r*-type effect size estimates using alternative statistical data (e.g., *t* or *F* values) with procedures described by Rosenthal (1991). If the authors reported effects as nonsignificant but did not provide any further statistical information, a conservative effect size of zero was assumed. For ease of expression, the term *effect size* is used henceforth, rather than the more precise term "*r*-type effect size estimate" (Rosenthal, Hoyt, Ferrin, Miller, & Cohen, 2006).

In order to preserve the independence of effect sizes, we included only one effect size from each subject sample in the analyses (Lipsey & Wilson, 2001). There were five cases in which a study reported results obtained with the same subject sample as another study within the current meta-analysis. (These studies are separated by dotted, rather than solid, lines in the online Supplementary Tables 1 and 2.) The effect sizes associated with these studies were averaged to compute an overall adherence–outcome or competence–outcome effect size for each subject sample.

In addition, some studies included more than one adherence– outcome or competence– outcome effect size. In most cases, this was because the authors reported correlations of adherence or competence measures with more than one outcome variable. For those studies in which a specific population (e.g., depressed patients) was targeted, we selected the effect size with the most relevant outcome measure (e.g., a measure of depressive symptoms). In contrast, for those studies that did not target a specific population or that did target a specific population but included more than one outcome measure to assess the target problem (e.g., Hall, 2007), the effect sizes reported were averaged to compute a single mean effect size for that study. However, in such cases, if one of the adherence– or competence–outcome effect sizes was based on an outcome measure assessed by an independent evaluator (e.g., Hamilton Rating Scale for Depression [HRSD]; Hamilton, 1960), this effect size was selected in preference to therapist- or patient-rated measures, as such evaluators generally are specifically trained to administer these measures and are typically blind to at least some sources of information that may bias such ratings (e.g., treatment group membership,

---

[1]Three effect sizes (the adherence–outcome finding from Shaw et al., 1999, and both the cognitive–behavioral therapy and interpersonal psychotherapy adherence–outcome findings from Elkin, 1988) were excluded from the analyses, as the direction of these effects were not ascertainable either from the manuscript or after contacting the authors directly.

treatment integrity, the outcomes of a given therapist's previous cases). If the choice was between a patient's self-report measure or a therapist's rating of the patient's improvement, only the effect size derived from the former measure was included.

Similarly, for those studies that reported several effect sizes resulting from correlating multiple measures of adherence or competence with outcome, these effect sizes were averaged. However, in several studies, adherence–outcome or competence–outcome relations were examined with subscales of adherence or competence measures assessing techniques that are considered to be particularly important within the treatment modality in question (i.e., Cognitive Therapy–Concrete subscale of the CSPRS for DeRubeis & Feeley, 1990; Feeley et al., 1999; Strunk, Brotman, & DeRubeis, 2009; Exploratory Intervention subscale of the Inventory of Therapeutic Strategies [ITS] for Gaston, Piper, Debbane, Bienvenu, & Garant, 1994; Gaston & Ring, 1992; Gaston, Thompson, Gallagher, Cournoyer, & Gagnon, 1998; Expressive Techniques subscale of the Penn Adherence/ Competence Scale for Supportive–Expressive Dynamic Psychotherapy [PAC–SE] for Barber et al., 1996). In these cases, the effect size based on these more focused adherence or competence subscales was selected. In addition to the adherence–outcome or competence– outcome effect sizes, each study was also coded for several possible moderators of these relations: (a) treatment modality (i.e., cognitive–behavioral therapy, interpersonal therapy, dynamic therapy, client-centered therapy, emotion-focused trauma therapy, brief relational therapy, individual drug counseling); (b) type of problem targeted (i.e., major depressive disorder, panic disorder, bulimia nervosa, bereavement, substance use, child abuse trauma, mix of diagnoses); (c) temporal confound (i.e., whether the symptom change that preceded the assessment of adherence or competence had been accounted for, methodologically or statistically); and (d) alliance confound (i.e., whether the influence of the therapeutic alliance had been controlled).

Effect sizes were weighted by the inverse of their variance, thus giving more weight to effect sizes from studies with larger samples (Rosenthal, 1991). To test whether mean weighted effect sizes significantly differed from zero, we calculated z tests (Lipsey & Wilson, 2001). In order to test whether the various effect sizes in this meta-analysis estimated the same population mean, we conducted a homogeneity analysis based on Hedges and Olkin's (1985) Q statistic. The Q statistic is designed to test whether the observed variability across effect sizes is greater than expected from subject-level sampling error. Potential moderators of effect size variability were then examined with a meta- analytic analogue to the analysis of variance (ANOVA; Lipsey & Wilson, 2001). For moderator analyses, groups consisting of only one effect size were excluded.

Given that the sampling distribution of rs is known to be skewed, particularly when the population value of r is large, all calculations and analyses were carried out on Fisher's Z values (Rosenthal, 1991). For the purposes of simplifying the presentation of results, however, once all analyses were conducted, Fisher's Z values were converted back to rs. We conducted all analyses using inverse-variance weighted, maximum likelihood, random effects models (Lipsey & Wilson, 2001). Our decision to use a random effects, rather than a fixed effects, model was informed by both conceptual and statistical considerations. Briefly, from a conceptual standpoint, the apparent differences across the included studies (e.g., in

treatment modality, problems targeted, methodology, and so on) led us to believe that the various effect sizes were not estimating a common population mean. In addition, the results from the homogeneity analyses, presented in the following section, challenged the assumptions of a fixed effects model. Comprehensive Meta-Analysis, Version 2.0, was used for all analyses (Borenstein, Hedges, Higgins, & Rothstein, 2005).

Results

Distribution of Effect Sizes

Supplementary Tables 1 and 2 (see online) present effect sizes and descriptive characteristics for all adherence and competence studies, respectively. In all cases, positive effect sizes indicate that higher adherence or competence ratings were associated with better outcomes. Adherence–outcome effects sizes ranged from −.40 to .47. The mean weighted adherence–outcome effect size was .02, which was not significantly different from 0 at the .05 level, $z = 0.36$; 95% confidence interval (CI) [−.069, .100]. Similarly, competence–outcome effect sizes ranged from −.36 to .73. The mean weighted competence-outcome effect size was .07, which was not significantly different from 0, $z = 0.97$, 95% CI [−.069, .201].

Heterogeneity of Effect Sizes

Figures 1 and 2 display forest plots of all effect sizes from the adherence and competence studies, respectively. Effect size distributions were significantly heterogeneous for both adherence– outcome ($Q = 50.90$, $p < .01$) and competence–outcome ($Q = 37.15$, $p < .01$) relations, suggesting that the various effect sizes were not estimating a common population mean in that their variability is greater than expected from sampling error alone.

Another value to consider in quantifying the heterogeneity of effect size distributions is I^2 (Higgins, Thompson, Deeks, & Altman, 2003), which reflects the percentage of total variation across effect sizes due to true heterogeneity (i.e., between-studies variability) rather than sampling error. Higgins et al. (2003) proposed benchmarks for I^2 whereby a value of 25% is considered low, 50% is moderate, and 75% is high. Moderate to high I^2 values were associated with both the adherence–outcome (47.0%) and the competence– outcome (59.6%) effect sizes, indicating that approximately half of the variability in effect sizes across both the adherence and competence studies was due to differences related to aspects of these studies, rather than sampling error. Potential moderators that might explain this variability were thus examined.

Moderator Analyses

Treatment modality—To examine whether treatment modality moderated adherence– outcome and competence–outcome effect sizes, we computed a mean weighted effect size for each of the treatment modalities included in this meta-analyses. With respect to adherence–outcome findings, mean weighted effect sizes were not significantly different across treatment modalities, $Q(3) = 2.61$, $p = .46$ (see Table 1).[2] Similarly, treatment modality did not emerge as a significant moderator of competence–outcome effect sizes, $Q(2) = 4.22$, $p = .12$ (see Table 2).

Problem targeted—Effect sizes were separated according to the primary mental health problem targeted in the studies, and a mean weighted effect size was computed for each of these groups. With respect to the adherence–outcome findings, mean weighted effect sizes were not significantly different across the types of problems targeted, $Q(4) = 4.02$, $p = .40$ (see Table 1). Although none of the mean weighted adherence–outcome effect sizes for the target problem groups were significantly different from zero, a nonsignificant trend emerged for the subset of studies that targeted clinical depression ($r = .12$; $z = 1.73$, $p = .08$). Among the competence–outcome findings, mean weighted effect sizes were significantly different across the types of problems targeted, $Q(3) = 18.33$, $p < .001$. As shown in Table 2, the largest mean weighted effect size was associated with those studies targeting clinical depression ($r = .28$).

Temporal confound—To examine whether effect sizes varied as a function of whether temporal confounds were controlled (i.e., whether subsequent symptom change was being predicted), we separated studies into the two relevant groups. Of the 26 studies in which the relation between adherence and outcome was examined, only eight controlled for temporal confounds.

The Feeley et al. (1999) and Castonguay et al. (1996) studies were based on the same subject sample. Thus, as noted earlier, in order to preserve the independence of effect sizes, in our primary analyses, we averaged the adherence–outcome findings from these two studies to form a single mean effect size. However, given that Feeley et al. controlled for temporal confounds but Castonguay et al. did not, we conducted two moderator analyses. In the first, the assumption of independence was relaxed, and the adherence–outcome effect sizes associated with each of the two studies were included in the analysis. There was not a significant difference in mean weighted effect sizes between those studies that adequately controlled ($r = .06$) and those studies that did not control ($r = -.01$) for temporal confounds, $Q(1) = 0.46$, $p = .50$ (see Table 1). Similarly, when the assumption of independence was upheld and these two studies were excluded from the moderator analyses, there was not a significant difference in mean weighted effect size between those studies that controlled ($r = .03$) and those that did not control ($r = .01$) for temporal confounds, $Q(1) = 0.01$, $p = .91$.

Although Barber et al. (2006, 2008) controlled for temporal confounds, we performed the latter two analyses again, excluding the effect sizes from these studies, as there is some reason to believe that the effect sizes they reported were overly conservative (see the Appendix). Indeed, when these two effect sizes were excluded, the mean weighted effect size for studies that controlled for temporal precedence increased to .16 (when the assumption of independence was relaxed) and .13 (when the assumption was upheld). However, the difference in effect size between those studies that adequately controlled and those studies that did not control for temporal confounds failed to reach statistical significance when the assumption of independence was relaxed, $Q(1) = 2.49$, $p = .12$, and when it was fully upheld, $Q(1) = .94$, $p = .33$.

---

[2]The Luborsky, McLellan, Woody, O'Brien, and Auerbach, (1985) study was excluded from this particular moderator analysis due to the fact that the adherence–outcome correlation reported in that study was based on a unique measure of therapist adherence (i.e., *purity*) that involved collapsing across three different treatment modalities. The Patton, Muran, Safran, Wachtel, and Winston (2009) study was also excluded from this analysis for a similar reason.

Of the 15 studies in which the relation between competence and outcome was examined, only four controlled for temporal confounds. There was not a significant difference in mean weighted effect sizes between those competence studies that adequately controlled ($r = -.03$) and those that did not control ($r = .11$) for temporal confounds, $Q(1) = 0.98$, $p = .32$ (see Table 2). Similarly, when the competence–outcome effect sizes from the Barber et al. (2006, 2008) studies were excluded, the mean weighted effect size for those studies that adequately controlled ($r = .09$) and those studies that did not control ($r = .11$) for temporal confounds were not significantly different, $Q(1) = 0.01$, $p = .95$.

Alliance confound—Eleven studies examined the relation between adherence and outcome while the influence of the therapeutic alliance was controlled. Their mean weighted effect size ($r = .08$) was compared with that of the 15 studies in which the alliance was not controlled ($r = .02$). The difference between these two effect sizes was not significant $Q(1) = 0.35$, $p = .55$ (see Table 1).

Of the 15 studies that examined the relation between competence and outcome, nine included a statistical control for the influence of the alliance. We compared the two sets of studies in two moderator analyses, one in which the assumption of independence was relaxed, and one in which it was upheld.[3] In both analyses, the mean weighted competence–outcome effect size was significantly smaller when the alliance was statistically controlled (rs = $-.03$ and $.00$) in comparison to when it was not (rs = $.23$ and $.26$), $Q(1) = 4.52$, $p < .05$, in the former analysis, and $Q(1) = 4.58$, $p < .05$, in the analysis in which the assumption of independence was upheld (see Table 2).

Within-study comparison—A within-study analysis was also conducted to examine the degree to which statistically controlling for the alliance may influence adherence– outcome and competence– outcome effect sizes. Nine studies examined adherence–outcome relations with and without statistical controls for the influence of the alliance. For each of these studies, we computed difference scores by subtracting the effect size obtained when the alliance was included as a covariate from the effect size when the alliance was not included in the analysis. The mean of these difference scores was $-.04$, which was not significantly different from 0 at the .05 level, $z = -0.67$; 95% CI [$-.140$, $.069$].

We conducted a parallel set of calculations with the seven studies that examined competence–outcome relations with and without statistical controls for the alliance. Similar to the results discussed earlier, the mean of these difference scores was $.01$, which was not significantly different from 0, $z = 0.16$; 95% CI [$-.098$, $.116$].[4]

---

[3]The two Svartberg and Stiles (1992; 1994) studies included in this meta-analysis were based on the same subject sample, and thus their competence–outcome effect sizes were pooled in the primary analyses (see supplementary Table 2). However, in only one of these two studies (Svartberg & Stiles, 1994) was the relation between competence and outcome examined while the influence of the alliance was controlled. Thus, when the assumption of independence was relaxed, the effect sizes associated with these studies were included in the moderator analysis. In contrast, when the assumption of independence was upheld, they were excluded.
[4]Similar to all other analyses, these calculations were carried out on Fisher's Z, rather than r (Rosenthal, 1991). For the purposes of simplifying the presentation of results, however, once all analyses were conducted, Fisher's Z values were converted back to rs.

Discussion

In this systematic review, we analyzed findings from 36 studies in which therapist adherence or competence was examined in relation to outcome. The most striking result is that variability in neither adherence nor competence was found to be related to patient outcome and indeed that the aggregate estimates of their effects were very close to zero. One explanation for these results is that adherence and competence are relatively inert therapeutic ingredients that play at most a small role in determining the extent of symptom change. It is possible that the constituent studies in which significant positive adherence–outcome and competence– outcome effect sizes were reported were simply chance findings from a population in which variability in adherence and competence accounts for little, if any, variability in outcome.

However, given that significant heterogeneity was observed across both the adherence–outcome and competence–outcome effect sizes, mean effect sizes must be interpreted with caution. Four plausible moderators of this variability were examined, two of which yielded significant results in the competence, but not the adherence, analyses. Specifically, there were indications that some of the heterogeneity across effect sizes could be due to differences in study methods. In particular, competence–outcome studies that controlled for the influence of the alliance reported significantly smaller effect sizes in comparison to those that did not, although this finding did not emerge in our within-study analysis. In addition, the type of patient's problem targeted emerged as a significant moderator of competence–outcome effect sizes, with the highest positive relations observed in studies of psychotherapy for major depressive disorder (MDD). Similarly, although the moderator effect was not significant in the adherence–outcome analysis, a nonsignificant trend emerged for the subset of studies that targeted MDD.

It may be that the symptoms of clinical depression, in contrast to the symptoms of the other disorders and mental health problems targeted in the included studies, are relatively more responsive to therapist interventions (at least those employed and measured in the treatment modalities represented in this meta-analysis). It is also important to note that, with the exception of the "mix of diagnoses" group for the adherence studies (see Table 1), there were more adherence and competence studies targeting MDD than any other disorder, resulting in greater power to observe effects in this group of studies.

If therapist adherence and competence play little or no role in producing symptom change, it may be that a more important set of factors are those that are common to most or all forms of psychotherapy, such as the quality of the therapeutic alliance. Indeed, meta-analyses of the literature on alliance–outcome relations (Horvath & Symonds, 1991; Martin et al., 2000) have reported larger average effect sizes than those we obtained in the present review. To maximize symptom change, it may be more important to focus on enhancing the dose of certain common factors such as the alliance, rather than increasing therapist adherence or competence. However, it is important to note that mean alliance– outcome correlations of . 26 (Horvath & Symonds, 1991) and .22 (Martin et al., 2000) are considered relatively small, accounting for only 7% and 5% of the variance in treatment outcome, respectively.

As discussed by DeRubeis and Webb (2010), several features common to studies of therapy process can limit the magnitude of estimates of process–outcome relations, and some of these may have been at play in the literature represented in the present meta-analysis. These include unreliability of measures and the possibility of restrictions in the ranges of the outcome variables and measured adherence and competence. Indeed, many of the studies included in our meta-analysis used data from investigations of therapy in which great care was taken to select, train, and monitor the delivery of therapy, especially in regard to adherence to, and competence in, the therapy under investigation. To the extent that this was the case, adherence– outcome and competence–outcome correlations may have been attenuated.

Therapist *responsiveness* may also help account for the failure of many authors to find significant, positive relations between outcome and seemingly important therapist variables such as adherence and competence. Responsiveness refers to the fact that therapists generally do not deliver predetermined levels of particular interventions (i.e., *ballistic* action) but rather adapt their behavior to the emerging context, in particular, patient behaviors (Stiles, Honos-Webb, & Surko, 1998). For example, a therapist might adhere more to the methods of a particular intervention with patients who are not improving and who appear to be at high risk for evidencing a relatively poor outcome. The presence of enough such cases in a study would be reflected in small or even negative correlations between method and outcome, even if experimental studies of these same phenomena would show that higher "doses" of the therapy, in the form of greater adherence or competence, lead to better outcomes, on average.

Furthermore, for the majority of the studies included in this review, adherence and competence scores were based on ratings of one or a few therapy sessions, and these were correlated with scores on an outcome measure obtained several weeks or months later. Causal relations may be obscured by processes that occur in the intervening time period. Indeed, Strunk et al. (2009) found that adherence was significantly related to short-term, but not longer term, outcome, in the same set of therapist–patient dyads. Specifically, they reported that ratings of adherence to concrete cognitive therapy techniques in early sessions predicted depressive symptom change from one session to the next, yet adherence ratings averaged across the first four sessions were not significantly related to symptom change from the fourth session to the end of treatment.

Implicit within the research designs employed in such studies is the assumption that ratings of adherence and competence represent the level of these variables that patients receive throughout therapy. To the extent that this is the case, process researchers would gain little, other than improving the reliability of such pooled scores, from assessing and averaging adherence and competence ratings over the course of multiple sessions. A single-session "snapshot" may be sufficient to accurately capture these constructs. However, if adherence and competence are relatively unstable over the course of therapy (i.e., low test–retest reliability), ratings based on only one or a few early sessions would yield unreliable estimates and, consequently, would not be expected to correlate highly with posttreatment outcome. Although the cross-session reliabilities of adherence and competence measures are not often reported, in two studies that were included in the present review, adherence and

competence ratings were found to vary substantially across sessions (Barber et al., 2006, 2008). Insofar as relationship variables such as the therapeutic alliance evidence greater stability over the course of treatment, such variables may be expected to correlate more highly with posttreatment outcome than adherence or competence, when ratings of these constructs are based on only one or a few sessions.

Estimates of the prediction of outcome from process variables may be biased when temporal confounds are not addressed in research designs and data analyses. As noted by Feeley et al. (1999), temporal confounds have often been ignored in studies of alliance–outcome relations. When temporal confounds have been controlled, alliance–outcome findings have been less consistent across studies (Barber, 2009; Strunk et al., 2009). The majority of the adherence and competence studies included in this meta-analysis also failed to control for temporal confounds. However, whereas it appears that temporal confounds may result in inflated estimates of alliance–outcome associations, we did not observe this pattern in our analyses of adherence–outcome or competence–outcome relations. Although speculative, one explanation could be that when the influence of prior symptom change is ignored, alliance–outcome relations are inflated, whereas adherence–outcome and competence–outcome relations are relatively unaffected or perhaps attenuated. One can imagine how early symptom change may contribute to an improvement in the alliance (Barber, Connolly, Crits-Cristoph, Gladis, & Siqueland, 2000; Feeley et al., 1999; Stiles, Shapiro, & Elliot, 1986). In contrast, it is less clear what impact early symptom change may have on therapist adherence. For example, in some cases, early symptom improvement may encourage therapists to deliver more, rather than fewer, interventions from the given therapy protocol (Loeb et al., 2005). In other cases, and as noted earlier, therapists may deliver more interventions to those patients whom they perceive as deteriorating or evidencing relatively little symptom improvement.

Furthermore, findings from a few of the studies reviewed (e.g., Barber et al, 2006, 2008; Hogue et al., 2008; Piper et al., 1991) suggest that, at least in some contexts, the relation between adherence or competence and outcome may be nonlinear (e.g., quadratic). In most of the studies included in this meta-analysis, only linear adherence– or competence–outcome relations were tested; therefore, underlying nonlinear effects may have gone undetected. Consequently, the current meta-analysis was restricted to a quantitative synthesis of linear adherence–outcome and competence–outcome relations. To the extent that the relation between adherence or competence and outcome is in fact nonlinear, this meta-analysis would represent an inappropriate modeling of actual curvilinear effects.

Finally, the focus in most of the included studies was on the association between scales assessing adherence to, or competence in, a variety of interventions or methods and outcome. It may be that only some of these interventions or methods are significantly associated with symptom change but that their relation with outcome is masked due to their being pooled with other interventions or methods that are relatively unrelated to outcome. For example, DeRubeis and colleagues (DeRubeis & Feeley, 1990; Feeley et al., 1999) have found that variability in therapist adherence to concrete, problem-focused cognitive therapy techniques predicted subsequent symptom change. In contrast, therapist adherence to more abstract, less focused cognitive therapy methods (e.g., discussing the relationship between

thoughts and feelings) has not been found to be significantly associated with symptom change in those same investigations. Furthermore, therapist interventions do not operate in a vacuum. Namely, the strength of the relation between particular therapist interventions or methods and patient outcome is likely moderated by a number of other therapist and patient variables, as well as interactions among these variables (e.g., patient motivation, the quality of the alliance). Despite the fact that this point is widely acknowledged by process researchers, in many of the studies we reviewed, adherence–outcome or competence–outcome relations were examined in isolation from other likely important process variables. More informative findings are likely to emerge from studies examining the joint contribution of, and interaction between, multiple process variables (both common and specific), across several time points (Barber, 2009). Such investigations would likely provide a more accurate picture of how these variables change and interact with one another over time to account for variability in outcome and may help move the field beyond the sometimes overly simplistic and dichotomous common factors versus specific factors debate.

Limitations

Several limitations of the current meta-analytic review should be noted. First, the relatively small number of adherence–outcome and competence–outcome effect sizes included in this meta-analysis limited power, particularly for the purposes of moderator analyses. Similarly, more precise I^2 values would have been obtained if more effects sizes had been included. Second, in part to increase power, we also included studies in which only particular components of adherence or competence (e.g., therapist use of interpretations in dynamic therapy) were assessed. Third, a relatively large number of moderator analyses were conducted, increasing the risk of Type 1 errors.

Future Directions

Although there is evidence in support of the efficacy of certain psychotherapies, the mechanisms through which these treatments exert their beneficial effects are generally not well understood (Kazdin, 2006). Undoubtedly, more studies need to be conducted, across a variety of different treatment modalities, in which process variables, such as therapist adherence and competence, are examined in relation to outcome. Additional studies investigating adherence–outcome and competence–outcome relations would also provide for more statistically powerful meta-analytic reviews.

Some recommendations for future process research can be derived from this meta-analysis. First, all of the studies included used observational, rather than experimental, methods. Although the statistical relation between adherence (or competence) and outcome was examined in all studies, very few of these studies controlled for both temporal confounds and plausible third variables. Results indicated that the alliance might be a third variable that should be controlled in future studies of adherence–outcome or competence–outcome relations. Furthermore, as noted by Barber (2009), more studies examining Alliance × Intervention interactions are also needed, as well as studies in which these relations are modeled as both linear and curvilinear effects. Of course, given the observational nature of these studies, even if plausible third variables, such as the alliance, are statistically

controlled, nonspuriousness cannot be ruled out. Namely, there may be one or more unmeasured third variables that account for the statistical relation between process variables and outcome.

In an effort to maximize internal validity and thus strengthen causal inferences, researchers should conduct more experimental studies of process–outcome relations, in which, ideally, one therapy process variable is manipulated, while all others are held constant and patients are randomly assigned to conditions (e.g., Høglend et al., 2008). Such methodologically rigorous studies will likely help researchers to isolate the active ingredients in various psychotherapies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Appendix

Additional Notes on Included Studies

Barber et al. (2006, 2008): The adherence and competence ratings from these two studies were based on randomly selected sessions between Sessions 2–10, while drug usage (outcome) was assessed at monthly intervals. Thus, outcome ratings were not always available at the same time point at which adherence and competence were assessed. As a result, for a number of patients, ratings of adherence or competence were correlated with symptom change beginning at some point after process measures were assessed. Such a data-analytic approach may yield overly conservative adherence–outcome or competence–outcome effect sizes as it does not capture the short-term effect of the latter process variables on outcome (Strunk et al. 2009).

Castonguay et al. (1996) and Feeley et al. (1999): The process–outcome correlations in these two studies were based on patients who received either cognitive therapy alone or cognitive therapy plus imipramine pharmacotherapy.

Gaston et al. (1998): This study also included a behavior therapy and cognitive therapy condition. However, given that the adherence measure (Inventory of Therapeutic Strategies; Gaston et al., 1994; Gaston et al., 1998; Gaston & Ring, 1992) employed was developed from a psychodynamic perspective, only the adherence–outcome effect size from the psychodynamic condition was included.

Kuyken and Tsivrikos (2009): Evaluation of Therapist's Behavior Form (ETBF) competence ratings were completed by Cory Newman, the Clinical Director at the Center for Cognitive Therapy. In their article, Kuyken and Tsivrikos (2009) stated that the director's ratings "were made on the basis of his knowledge of the center therapists' work as a whole, taking into account all available information" (p. 44). According to Kuyken, for the majority of therapists, ETBF ratings were based in part on audiotape recordings of therapy sessions (W. Kuyken, personal communication, January 5, 2009).

Minonne (2008): The author conducted two sets of analyses, controlling for pretreatment Beck Depression Inventory (BDI; Beck, Steer, and Brown, 1996) and Vanderbilt Therapeutic Alliance Scale (Hartley & Strupp, 1983) in both: one predicting residual change in depression scores from mean process ratings at Sessions 4, 7 or 8, and 14 or 15, and one predicting residual change in depression from mean process ratings at Sessions 7 or 8 and 14 or 15. Effect sizes from these two analyses were averaged.

Strunk et al. (2009): The adherence–outcome effect size extracted from this study was derived from a repeated-measure regression analysis examining the relation between adherence ratings at Sessions 1–4 and BDI ratings at Session $n + 1$ (with BDI at Session n and site included as covariates). A secondary analysis conducted by Strunk et al., which tested the relation between adherence ratings in the first four sessions (averaged) and the rate of improvement following Session 4 (through the end of treatment), was not included in the current meta-analysis as it was not the author's primary hypothesis.

References

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

Barber JP. Toward a working through of some core conflicts in psychotherapy research. Psychotherapy Research. 2009; 19:1–12. [PubMed: 19206018]

Barber JP, Connolly MB, Crits-Cristoph P, Gladis L, Siqueland L. Alliance predicts patients' outcome beyond in-treatment change in symptoms. Journal of Consulting and Clinical Psychology. 2000; 68:1027–1032. [PubMed: 11142536]

*Barber JP, Crits-Christoph P, Luborsky L. Effects of therapist adherence and competence on patient outcome in brief dynamic therapy. Journal of Consulting and Clinical Psychology. 1996; 64:619–622. [PubMed: 8698958]

*Barber JP, Gallop R, Crits-Christoph P, Barrett MS, Klostermann S, McCarthy KS, Sharpless BA. The role of the alliance and techniques in predicting outcome of supportive–expressive dynamic therapy for cocaine dependence. Psychoanalytic Psychology. 2008; 25:461–482.

*Barber JP, Gallop R, Crits-Christoph P, Frank A, Thase ME, Weiss RD, Gibbons MB. The role of therapist adherence, therapist competence, and the alliance in predicting outcome of individual drug counseling: Results from the NIDA Collaborative Cocaine Treatment Study. Psychotherapy Research. 2006; 16:229–240.

Beck, AT.; Steer, RA.; Brown, GK. Manual for the Beck Depression Inventory. 2nd ed.. Psychological Corporation; San Antonio, TX: 1996.

Borenstein, M.; Hedges, L.; Higgins, J.; Rothstein, H. Comprehensive Meta-Analysis, Version 2 [Computer software]. Biostat; Englewood, NJ: 2005.

*Carroll KM, Nich C, Rounsaville BJ. Contribution of the therapeutic alliance to outcome in active versus control psychotherapies. Journal of Consulting and Clinical Psychology. 1997; 65:510–514. [PubMed: 9170775]

Castonguay LG. "Common factors" and "nonspecific variables": Clarification of the two concepts and recommendations for research. Journal of Psychotherapy Integration. 1993; 3:267–286.

*Castonguay LG, Goldfried MR, Wiser S, Raue PJ, Hayes AM. Predicting the effect of cognitive therapy for depression: A study of unique and common factors. Journal of Consulting and Clinical Psychology. 1996; 64:497–504. [PubMed: 8698942]

Castonguay LG, Holtforth MG. Change in psychotherapy: A plea for no more "nonspecific" and false dichotomy. Clinical Psychology: Science and Practice. 2005; 12:198–201.

Chambless DL, Ollendick TH. Empirically supported psychological interventions: Controversies and evidence. Annual Review of Psychology. 2001; 52:685–716.

*Crits-Christoph P, Cooper A, Luborsky L. The accuracy of therapists' interpretations and the outcome of dynamic psychotherapy. Journal of Consulting and Clinical Psychology. 1988; 56:490–495. [PubMed: 3198804]

*DeRubeis RJ, Feeley M. Determinants of change in cognitive therapy for depression. Cognitive Therapy and Research. 1990; 14:469–482.

DeRubeis, RJ.; Webb, CA. The attenuation of process–outcome relations. 2010. Manuscript in preparation

DeRubeis, RJ.; Webb, CA.; Tang, TZ.; Beck, AT. Cognitive therapy.. In: Dobson, KS., editor. Handbook of cognitive–behavioral therapies. 3rd ed.. Guilford; New York, NY: 2009.

Dimidjian S, Hollon SD, Dobson KS, Schmaling KB, Kohlenberg RJ, Addis M, Dunner DL. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. Journal of Consulting and Clinical Psychology. 2006; 74:658–670. [PubMed: 16881773]

*Elkin, I. Relationship of therapists' adherence to treatment outcome in the Treatment of Depression Collaborative Research Program.. Paper presented at the meeting of the Society for Psychotherapy Research in Santa Fe; New Mexico. 1988, June;

Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM, Collins JF, Parloff MB. National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. Archives of General Psychiatry. 1989; 46:971–982. [PubMed: 2684085]

Elvins R, Green J. The conceptualization and measurement of therapeutic alliance: An empirical review. Clinical Psychology Review. 2008; 28:1167–1187. [PubMed: 18538907]

*Feeley M, DeRubeis RJ, Gelfand L. The temporal relation of adherence and alliance to symptom change in cognitive therapy for depression. Journal of Consulting and Clinical Psychology. 1999; 67:578–582. [PubMed: 10450629]

*Gaston L, Piper WE, Debbane EG, Bienvenu JP, Garant J. Alliance and technique for predicting outcome in short and long term analytic psychotherapy. Psychotherapy Research. 1994; 4:121–135.

*Gaston L, Ring JM. Preliminary results on the Inventory of Therapeutic Strategies (ITS). Journal of Psychotherapy Research and Practice. 1992; 1:1–13.

*Gaston L, Thompson L, Gallagher D, Cournoyer L, Gagnon R. Alliance, technique, and their interactions in predicting outcome of behavioral, cognitive, and brief dynamic therapy. Psychotherapy Research. 1998; 8:190–209.

Høglend P, Bøgwald KP, Amlo S, Marble A, Ulberg R, Sjaastad MC, Johansson P. Transference interpretations in dynamic psychotherapy: Do they really yield sustained effects? American Journal of Psychiatry. 2008; 165:763–771. [PubMed: 18413707]

*Hall, IE. Therapist adherence and technical skills in two versions of emotion-focused trauma therapy. University of Windsor; Windsor, Ontario, Canada: 2007. Unpublished doctoral dissertation

Hamilton MA. A rating scale for depression. Journal of Neurology, Neurosurgery, and Psychiatry. 1960; 23:56–61.

Hartley, DE.; Strupp, HH. The therapeutic alliance: Its relationship to outcome in brief psychotherapy.. In: Masling, J., editor. Empirical studies of psychoanalytical theories. Vol. 1. Erlbaum; Hillsdale, NJ: 1983.

Hedges, LV.; Olkin, I. Statistical methods for meta-analysis. Academic Press; Orlando, FL: 1985.

Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. British Medical Journal. 2003; 327:557–560. [PubMed: 12958120]

Hill CE, O'Grady KE, Elkin I. Applying the Collaborative Study Psychotherapy Rating Scale to rate therapist adherence in cognitive–behavior therapy, interpersonal therapy and clinical management. Journal of Consulting and Clinical Psychology. 1992; 60:73–79. [PubMed: 1556289]

*Hoffart A, Sexton H, Nordahl HM, Stiles TC. Connection between patient and therapist and therapist's competence in schema-focused therapy of personality problems. Psychotherapy Research. 2005; 15:409–441.

*Hogue A, Henderson CE, Dauber S, Barajas PS, Fried A, Liddle HA. Treatment adherence, competence, and outcome in individual and family therapy for adolescent behavior problems. Journal of Consulting and Clinical Psychology. 2008; 76:544–555. [PubMed: 18665684]

Hollon, SD.; Evans, MD.; Auerbach, A.; DeRubeis, RJ.; Elkin, I.; Lowery, A.; Piasecki, J. Development of a system for rating therapies for depression: Differentiating cognitive therapy, interpersonal psychotherapy, and clinical management pharmacotherapy. Vanderbilt University; Nashville, Tennessee: 1988. Unpublished manuscript

*Horowitz MJ, Marmar CR, Weiss DS, DeWin KN, Rosenbaum R. Brief psychotherapy of bereavement reactions: The relationship of process to outcome. Archives of General Psychiatry. 1984; 41:43–48.

Horvath AO, Symonds BD. Relation between working alliance and outcome in psychotherapy: A meta-analysis. Journal of Counseling Psychology. 1991; 38:139–149.

*Huppert JD, Barlow DH, Gorman JM, Shear MK, Woods SW. The interaction of motivation and therapist adherence predict outcome in cognitive behavioral therapy for panic disorder: Preliminary findings. Cognitive and Behavioural Practice. 2006; 13:198–204.

Judd CM, Kenny DA. Process analysis: Estimating mediation in treatment evaluations. Evaluation Review. 1981; 5:602–619.

Kazdin, AE. Mechanisms of change in psychotherapy: Advances, breakthroughs, and cutting-edge research (do not yet exist).. In: Bootzin, RR.; McKnight, PM., editors. Strengthening research methodology: Psychological measurement and evaluation. American Psychological Association; Washington, DC: 2006. p. 77-101.

Klonsky, ED. [December 9, 2009] Society of Clinical Psychology: American Psychological Association, Division 12. Website on research-supported psychological treatments. 2009. from Stony Brook University, Department of Psychology website: http://www.psychology.sunysb.edu/eklonsky-/division12/

*Kuyken W, Tsivrikos D. Therapist competence, co-morbidity and cognitive–behavioral therapy for depression. Psychotherapy and Psychosomatics. 2009; 78:42–48. [PubMed: 19018156]

Lipsey, MW.; Wilson, DB. Practical meta-analysis: Applied Social Research Methods Series. Vol. 49. Sage; Thousand Oaks, CA: 2001.

*Loeb KL, Wilson GT, Labouvie E, Pratt EM, Hayaki J, Walsh BT, Fairburn CG. Therapeutic alliance and treatment adherence in two interventions for bulimia nervosa: A study of process and outcome. Journal of Consulting and Clinical Psychology. 2005; 73:1097–1107. [PubMed: 16392983]

*Luborsky L, McLellan AT, Woody GE, O'Brien CP, Auerbach A. Therapist success and its determinants. Archives of General Psychiatry. 1985; 42:602–661. [PubMed: 4004503]

Martin DJ, Garske JP, Davis MK. Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. Journal of Consulting and Clinical Psychology. 2000; 68:438–450. [PubMed: 10883561]

*Marziali EA. Prediction of outcome of brief psychotherapy from therapist interpretive interventions. Archives of General Psychiatry. 1984; 41:301–304. [PubMed: 6703847]

*Minonne, GA. Therapist adherence, patient alliance, and depression change in the NIMH Treatment for Depression Collaborative Research Program (Doctoral dissertation). 2008. Available from ProQuest Dissertations and These database. (UMI No. 3328910)

Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA statement. Journal of Clinical Epidemiology. 2009; 62:1006–1012. [PubMed: 19631508]

*Ogrodniczuk JS, Piper WE. Measuring therapist technique in psychodynamic psychotherapies, development, and use of a new scale. Journal of Psychotherapy Practice and Research. 1999; 8:142–154. [PubMed: 10079461]

*Ogrodniczuk, JS. Therapist adherence to treatment manuals and its relation to the therapeutic alliance and therapy outcome: Scale development and validation (Doctoral dissertation). 1997. Available from Pro-Quest Dissertations and These database. (UMI No. NQ23049)

*O'Malley SS, Foley SH, Rounsaville BJ, Watkins JT, Sotsky SM, Imber SD, Elkin I. Therapist competence and patient outcome in interpersonal psychotherapy of depression. Journal of Consulting and Clinical Psychology. 1988; 56:496–501. [PubMed: 3198805]

*Paivio SC, Holowaty KAM, Hall IE. The influence of therapist adherence and competence on client reprocessing of child abuse memories. Psychotherapy. 2004; 41:56–68.

*Patton, J.; Muran, C.; Safran, JD.; Wachtel, P.; Winston, A. Measuring fidelity in three brief psychotherapies. 2009. Unpublished manuscript

Perepletchikova F, Kazdin AE. Treatment integrity and therapeutic change: Issues and research recommendations. Clinical Psychology: Science and Practice. 2005; 12:365–383.

*Piper WE, Azim HFA, Joyce AS, McCallum M. Transference interpretations, therapeutic alliance, and outcome in short-term individual psychotherapy. Archives of General Psychiatry. 1991; 48:946–953. [PubMed: 1929765]

*Piper WE, Debbane EG, Bienvenu JP, de Carufel F, Garant J. Relationships between the object focus of therapist interpretations and outcome in short-term individual psychotherapy. British Journal of Medical Psychology. 1986; 59:1–11. [PubMed: 3964579]

Rosenthal D, Hoyt W, Ferrin J, Miller S, Cohen N. Advanced methods in meta-analytic research: Applications and implications for rehabilitation counseling research. Rehabilitation Counseling Bulletin. 2006; 49:234–246.

Rosenthal, R. Meta-analytic procedures for social research. Sage; Newbury Park, CA: 1991.

*Sachs JS. Negative factors in brief psychotherapy: An empirical assessment. Journal of Consulting and Clinical Psychology. 1983; 51:557–564. [PubMed: 6619363]

Sharpless BA, Barber JP. A conceptual and empirical review of the meaning, measurement, development, and teaching of intervention competence in clinical psychology. Clinical Psychology Review. 2009; 29:47–56. [PubMed: 18952334]

*Shaw BF, Elkin I, Yamaguchi J, Olmsted M, Vallis TM, Dobson KS, Imber SD. Therapist competence ratings in relation to clinical outcome in cognitive therapy of depression. Journal of Consulting and Clinical Psychology. 1999; 67:837–846. [PubMed: 10596506]

*Spektor, D. Therapists' adherence to manualized treatments in the context of ruptures (Doctoral dissertation). 2008. Available from ProQuest Dissertations and These database. (UMI No. 3333897)

Stiles WB, Honos-Webb L, Surko M. Responsiveness in psychotherapy. Clinical Psychology: Science and Practice. 1998; 5:439–458.

Stiles WB, Shapiro DA, Elliott R. Are all psychotherapies equivalent? American Psychologist. 1986; 41:65–80.

*Strunk, DR.; Brotman, MA.; DeRubeis, RJ. The process of change in cognitive therapy for depression: Predictors of early intersession symptom gains and continued response. 2009. Manuscript submitted for publication

*Svartberg M, Stiles TC. Predicting patient change from therapist competence and patient–therapist complementarity in short-term anxiety-provoking psychotherapy: A pilot study. Journal of Consulting and Clinical Psychology. 1992; 60:304–307. [PubMed: 1592962]

*Svartberg M, Stiles TC. Therapeutic alliance, therapist competence, and client change in short-term anxiety-provoking psychotherapy. Psychotherapy Research. 1994; 4:20–33.

*Trepka C, Rees A, Shapiro DA, Hardy GE, Barkham M. Therapist competence and outcome of cognitive therapy for depression. Cognitive Therapy and Research. 2004; 28:143–157.

Wampold, BE. The great psychotherapy debate: Models, methods, and findings. Erlbaum; Mahwah, NJ: 2001.

*Webb, CA.; Gelfand, LA.; DeRubeis, RJ.; Amsterdam, JD.; Shelton, RC.; Hollon, SD.; Dimidjian, S. Mechanisms of change in cognitive therapy for depression: Therapist adherence, symptom change, and the mediating role of patient skills. 2009. Manuscript in preparation

Young, JE.; Beck, AT. The development of the Cognitive Therapy Scale. University of Pennsylvania; Philadelphia, PA: 1980. Unpublished manuscript

Study name	Statistics for each study					Correlation and 95% CI
	Correlation	Lower limit	Upper limit	Z-Value	p-Value	
Feeley et al. (1999)/Castonguay et al. (1996)	-0.03	-0.40	0.35	-0.15	0.88	
DeRubeis & Feeley (1990)	0.24	-0.17	0.58	1.15	0.25	
Webb et al. (2009)/Strunk et al. (2009)	0.34	0.12	0.53	3.01	0.00	
Minnone (2008) - CB	-0.07	-0.34	0.20	-0.52	0.60	
Huppert et al. (2006)	-0.24	-0.48	0.03	-1.75	0.08	
Hogue et al. (2008)	0.21	-0.04	0.44	1.68	0.09	
Loeb et al. (2005) - CB	-0.19	-0.49	0.15	-1.09	0.28	
Carroll et al. (1997)	0.00	-0.27	0.27	0.00	1.00	
Minnone (2008) - IP	0.11	-0.16	0.36	0.79	0.43	
Loeb et al. (2005) - IP	0.05	-0.27	0.36	0.30	0.77	
Barber et al. (1996)	0.10	-0.28	0.45	0.49	0.63	
Barber et al. (2008)	-0.21	-0.40	0.00	-1.94	0.05	
Luborsky et al. (1985)	0.47	0.19	0.68	3.14	0.00	
Patton et al. (2009)	0.10	-0.36	0.52	0.41	0.68	
Barber et al. (2006)	-0.09	-0.30	0.13	-0.77	0.44	
Marziali (1984)	0.31	-0.10	0.63	1.50	0.13	
Piper et al. (1986)/Gaston et al. (1994) -STD	-0.08	-0.52	0.39	-0.32	0.75	
Gaston et al. (1994) -LTD	0.20	-0.35	0.65	0.70	0.48	
Piper et al. (1991)	-0.34	-0.54	-0.10	-2.77	0.01	
Gaston et al. (1998)/Gaston and Ring (1992)	-0.09	-0.59	0.46	-0.30	0.76	
Horowitz et al. (1984)	0.00	-0.27	0.27	0.00	1.00	
Paivio et al. (2004)	0.09	-0.24	0.40	0.53	0.60	
Hall (2007) - IC	-0.28	-0.64	0.19	-1.19	0.24	
Hall (2007) - EE	-0.40	-0.68	-0.01	-2.03	0.04	
Ogrodniczuc (1997) - STI	0.01	-0.22	0.25	0.11	0.91	
Ogrodniczuk (1997) - SUP	0.05	-0.18	0.28	0.44	0.66	
Spektor (2008) - CB	0.06	-0.49	0.57	0.18	0.86	
Spektor (2008) - BR	0.32	-0.25	0.73	1.11	0.27	
	0.02	-0.07	0.10	0.36	0.72	

-1.00   -0.50   0.00   0.50   1.00

Figure 1.
Forest plot of all adherence–outcome effect sizes. R-type effect sizes are represented as squares, bounded by 95% confidence intervals (CIs), which are represented as lines. The area of each square is proportional to the inverse-variance weight assigned to each effect size. The diamond represents the mean weighted adherence–outcome effect size, and its width represents its 95% CIs. CB = Cognitive–behavioral; IP = interpersonal; STD = short-term dynamic; LTD = long-term dynamic; IC = imaginal confrontation; EE = evocative empathy; STI = short-term interpretive; SUP = short-term supportive; BR = brief relational.

Study name	Statistics for each study					Correlation and 95% CI
	Correlation	Lower limit	Upper limit	Z-Value	p-Value	
Shaw et al. (1999)	0.14	-0.20	0.45	0.81	0.42	
Kuyken & Tsivrikos (2009)	0.32	0.09	0.52	2.69	0.01	
Trepka et al. (2004)	0.28	-0.09	0.58	1.49	0.13	
Hogue et al. (2008)	-0.12	-0.36	0.13	-0.96	0.34	
Hoffart et al. (2005)	-0.04	-0.37	0.30	-0.23	0.82	
Barber et al. (1996)	0.25	-0.13	0.56	1.29	0.20	
Barber et al. (2008)	-0.23	-0.40	-0.04	-2.41	0.02	
Barber et al. (2006)	0.05	-0.17	0.26	0.41	0.68	
Crits-Christoph et al. (1988)	0.27	-0.06	0.54	1.63	0.10	
Svartberg & Stiles (1992; 1994)	-0.36	-0.75	0.22	-1.24	0.22	
Paivio et al. (2004)	-0.02	-0.34	0.31	-0.12	0.91	
Hall (2007) - IC	-0.26	-0.63	0.21	-1.10	0.27	
Hall (2007) - EE	-0.35	-0.65	0.04	-1.75	0.08	
Sachs (1983) - CC	0.45	-0.31	0.86	1.19	0.24	
Sachs (1983) - DYN	0.73	0.13	0.94	2.27	0.02	
O'Malley et al. (1988)	0.36	0.02	0.62	2.05	0.04	
	0.07	-0.07	0.20	0.97	0.33	

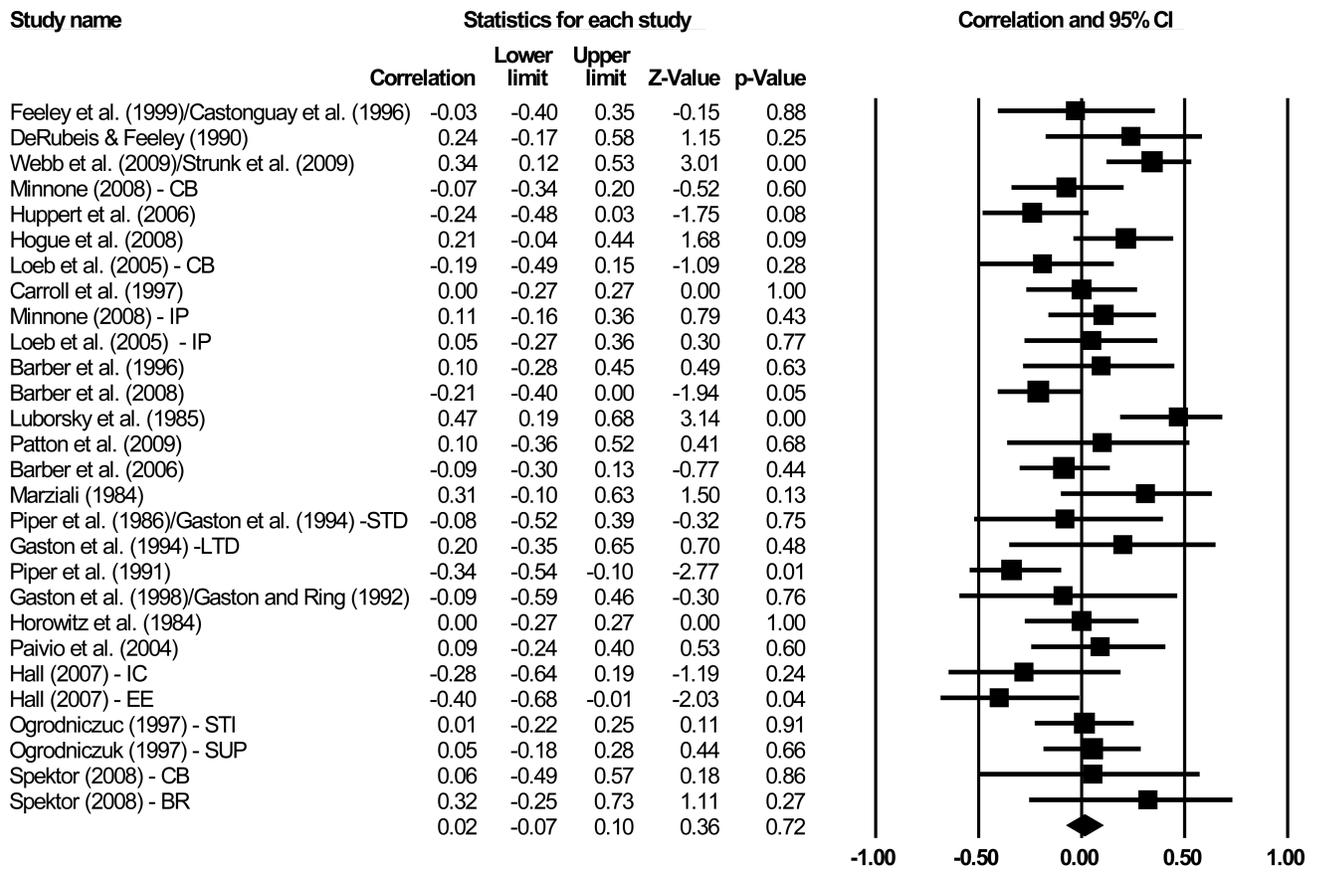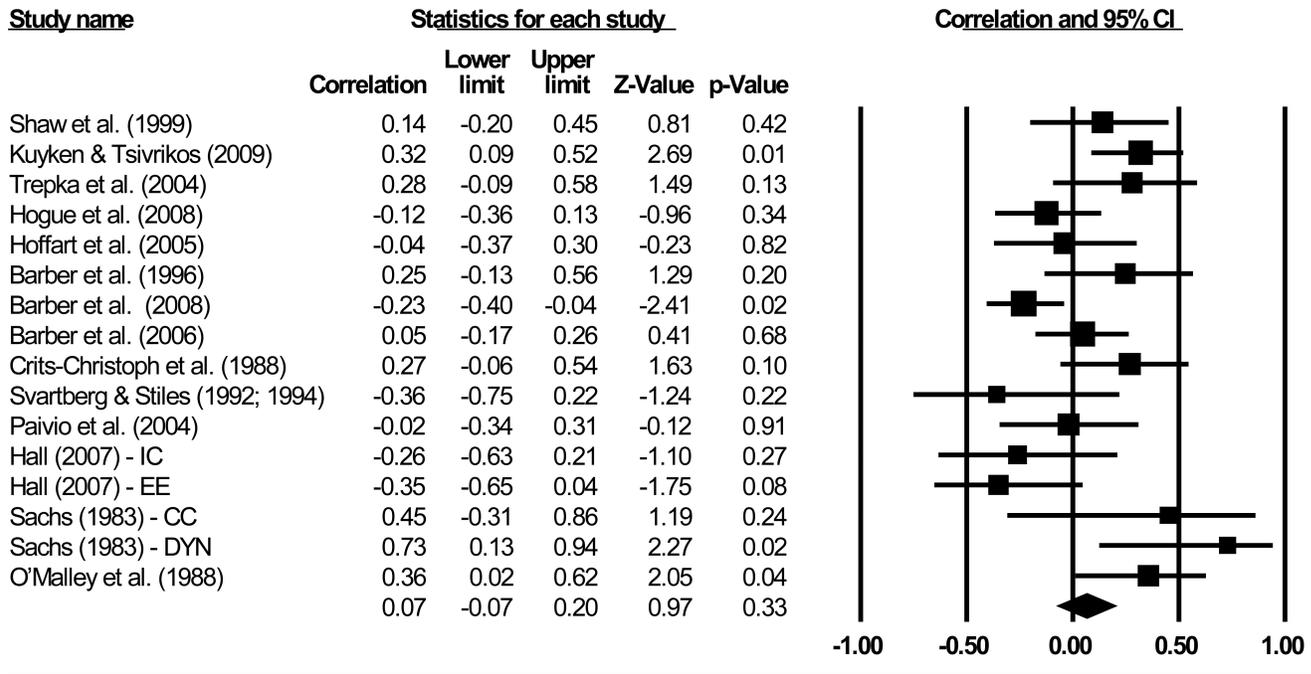-1.00 -0.50 0.00 0.50 1.00

Figure 2.
Forest plot of all competence–outcome effect sizes. R-type effect sizes are represented as squares, bounded by 95% confidence intervals (CIs), which are represented as lines. The area of each square is proportional to the inverse-variance weight assigned to each effect size. The diamond represents the mean weighted competence-outcome effect size, and its width represents its 95% CIs. IC = imaginal confrontation; EE = evocative empathy; CC = client-centered; DYN = dynamic.

Table 1

Moderator Analyses for Adherence-Outcome Effect Sizes

Moderator/subgroup	Mean effect size	SE	95% Confidence interval		Q value	df	I²
			Lower	**Upper**			
Treatment modality							
IP	.08	.105	−.12	.28	0.07	1	0.00
CB	.04	.077	−.11	.19	16.59*	8	51.77
Dynamic	−.04	.063	−.17	.08	13.29	9	32.27
EFTT	−.18	.162	−.47	.14	4.00	2	50.05
Problem targeted							
Depression	.12	.069	−.02	.25	7.29	6	17.64
Drug use	.07	.118	−.16	.29	16.86**	4	76.28
Mix of diagnoses	.02	.079	−.14	.17	12.11	8	58.36
Bulimia	−.07	.122	−.30	.17	0.98	1	0.00
Child abuse trauma	−.18	.162	−.47	.14	4.00	2	50.05
Temporal confound							
Controlled	.06	.091	−.11	.24	19.51**	7	64.12
Not controlled	−.01	.055	−.11	.10	40.67**	20	50.82
Alliance confound							
Controlled	.08	.084	−.09	.24	34.57**	10	71.07
Not controlled	.02	.056	−.09	.13	26.68*	16	40.03

Note. Positive effect sizes indicate that higher adherence ratings were associated with better outcomes. Significant Q values indicate that there is significantly more variability across effect sizes within a group than would be expected from subject-level sampling error. SE = standard error; CB = cognitive-behavioral; IP = interpersonal; EFTT = emotion-focused trauma therapy. df = degrees of freedom.

[*] $p < .05$.

[**] $p < .01$.

Table 2

Moderator Analyses for Competence-Outcome Effect Sizes

Moderator/subgroup	Mean effect size	SE	95% Confidence interval		Q value	df	I²
			Lower	Upper			
Treatment modality							
CB	.12	.098	−.08	.30	8.08	4	50.50
Dynamic	.11	.180	−.25	.43	16.59**	4	75.89
EFTT	−.18	.112	−.39	.04	1.79	2	0.00
Problem targeted							
Depression	.28**	.068	.14	.41	1.09	4	0.00
Mix of diagnoses	.27	.212	−.19	.63	7.37	3	59.29
Drug use	−.11	.086	−.27	.06	3.58	2	44.13
Child abuse trauma	−.18	.112	−.39	.04	1.79	2	0.00
Temporal confound							
Controlled	−.03	.103	−.23	.17	6.68	3	55.11
Not controlled	.11	.087	−.07	.27	26.02**	11	57.73
Alliance confound							
Controlled	−.03	.081	−.18	.13	21.33*	9	57.80
Not controlled	.23*	.087	.05	.39	8.31	6	27.82

Note. Positive effect sizes indicate that higher competence ratings were associated with better outcomes Significant *Q* values indicate that there is significantly more variability across effect sizes within a group than would be expected from subject-level sampling error. *SE* = standard error; CB = cognitive-behavioral; EFTT = emotion-focused trauma therapy. *df* = degrees of freedom.

[*] $p < .05$.

[**] $p < .01$.