Untangling the Alliance-Outcome Correlation: Exploring the Relative Importance of Therapist and Patient Variability in...

Article *in* Journal of Consulting and Clinical Psychology · January 2008

DOI: 10.1037/0022-006X.75.6.842 · Source: PubMed

CITATIONS
314

READS
1,927

3 authors, including:

Bruce Wampold
University of Wisconsin--Madison and Modu...
238 PUBLICATIONS **10,841** CITATIONS
SEE PROFILE

Zac E Imel
University of Utah
74 PUBLICATIONS **2,045** CITATIONS
SEE PROFILE

Journal of Consulting and Clinical Psychology
2007, Vol. 75, No. 6, 842–852

0022-006X/07/$12.00 DOI: 10.1037/0022-006X.75.6.842

Untangling the Alliance–Outcome Correlation: Exploring the Relative Importance of Therapist and Patient Variability in the Alliance

Scott A. Baldwin
Brigham Young University

Bruce E. Wampold and Zac E. Imel
University of Wisconsin—Madison

Although the therapeutic alliance is a consistent predictor of psychotherapy outcomes, research has not distinguished between the roles of patient and therapist variability in the alliance. Multilevel models were used to explore the relative importance of patient and therapist variability in the alliance as they relate to outcome among 331 patients seen by 80 therapists (therapist average caseload was 4.1). Patients rated both the alliance and outcome and all models adjusted for baseline psychological functioning. The results indicated that therapist variability in the alliance predicted outcome, whereas patient variability in the alliance was unrelated to outcome. Reasons why therapist variability as opposed to patient variability predicted outcome are discussed. Clinical implications include therapists monitoring their contribution to the alliance, clinics providing feedback to therapists about their alliances, and therapists receiving training to develop and maintain strong alliances.

Keywords: alliance, psychotherapy outcome, multilevel models, between-therapist correlation, within-therapist correlation

The effectiveness of psychotherapy is well established (Lambert & Bergin, 1994; Lambert & Ogles, 2004), although identifying the therapeutic factors that account for patient improvement has proved difficult (cf. Castonguay & Holtforth, 2005; Craighead, Bjornsson, & Amarson, 2005; DeRubeis, Brotman, & Gibbons, 2005; Kazdin, 2005; Wampold, 2005). One viable therapeutic factor is the therapeutic alliance, which has emerged as a consistent predictor of outcome (Horvath & Bedi, 2002; Horvath & Symonds, 1991; Martin, Garske, & Davis, 2000), leading many to conclude that it is an essential aspect of therapy (Norcross, 2002).

Alliance Theory and Research

The therapeutic alliance has its roots in psychodynamic theory, where it was defined as the healthy, trusting aspects of the patient–therapist relationship. Although there is presently no single agreed upon definition of the alliance, many researchers have adopted Bordin's (1979) pantheoretical definition of the alliance. As Hatcher and Barends (2006) pointed out, "The essence of Bordin's definition is this: Alliance describes the degree to which the therapy dyad is engaged in collaborative, purposive work" (p. 293).

Bordin's (1979) formulation of the alliance has three important implications (Hatcher & Barends, 2006). First, any aspect of intervention (e.g., assessment, case conceptualization, specific techniques) that relates to engagement in purposive work contributes to the alliance. Indeed, as Hatcher and Barends (2006) noted, ". . . alliance and technique occupy different conceptual levels and cannot be considered to be two different types of activity in therapy. Technique is an activity, alliance is a way to characterize activity" (p. 294). Second, the alliance is not equivalent to the therapeutic relationship, although the relationship may affect the alliance (i.e., the degree of collaboration and purposive work). Third, the alliance is not reducible to patients' experience, although patients' experience may provide a reasonable estimate of the alliance.

This pantheoretical conceptualization of the alliance allows researchers to study the relationship between the alliance and outcome within a number of therapies. In general, the alliance has been shown to be a consistent predictor of outcome across therapy approaches (Horvath & Bedi, 2002; Horvath & Symonds, 1991; Martin et al., 2000). In the most recent meta-analysis of the alliance–outcome correlation, A. E. Horvath and Bedi (2002) found that the average alliance–outcome correlation was .21, indicating that the alliance accounts for about 5% of the variance in outcome.

Interpreting the Alliance–Outcome Correlation

The alliance is formed in the complex transaction between the therapist and patient, each of whom brings to therapy his or her own characteristics, personality, and history (Gelso & Carter, 1994). The correlation between alliance and outcome does not account for this complexity, making it difficult to understand why the alliance is important to outcome. DeRubeis et al. (2005)

Scott A. Baldwin, Department of Psychology, Brigham Young University; Bruce E. Wampold and Zac E. Imel, Department of Counseling Psychology, University of Wisconsin—Madison.

This research was supported in part by a grant from the Graduate School of the University of Wisconsin—Madison awarded to Bruce E. Wampold. Portions of this research were presented at the annual meeting of the North American Society for Psychotherapy Research, Burr Oak, Ohio, October 2006. We thank Joseph Olsen for his time spent discussing the ideas and analyses presented in this article. Any errors that remain are our responsibility.

Correspondence concerning this article should be addressed to Scott A. Baldwin, 268 TLRB, Brigham Young University, Provo, UT 84602. E-mail: scott_baldwin@byu.edu

identified four sources of variability in the alliance, each of which may be related to outcome. The first source is related to the patient. Some patients may be better able to form a collaborative relationship than others. For example, patients' attachment style and social competencies may affect their ability to foster a strong alliance with their therapist (Mallinckrodt, 2000). The second source of variability in the alliance is related to the therapist. From this perspective, effective therapists may be able to engage patients in collaborative, purposive work, whereas ineffective therapists may be less able to do so. Thus, the alliance–outcome correlation would be due to variability in therapists' ability to form alliances with various patients. This perspective was proposed early on by Rogers (1957), who discussed therapists' capacity to be genuine and empathic and to show unconditional positive regard toward patients.

The third source of variability in the alliance is related to the interaction between patients and therapists. For example, some therapists may be able to form strong alliances with their patients regardless of their patients' abilities to form an alliance, whereas other therapists may be able to form a strong alliance only with those patients who come to therapy able to form strong alliances. Finally, the fourth source of variability in the alliance is related to the hypothesis that the alliance is a consequence of good outcomes. That is, change in therapy produces strong alliances, not the other way around.

Within- and Between-Therapist Correlations

The distinction between patient and therapist variability in the alliance dovetails nicely with the methodological distinction between within- and between-therapist correlations (Raudenbush & Bryk, 2002; Snijders & Bosker, 1999). Therapy data are hierarchical and include at least two levels: (a) the therapist level and (b) the patient level, where patients are nested within therapists. Consequently, we can consider relationships between variables at the therapist level (i.e., between-therapist correlations) and at the patient level (i.e., within-therapist correlations). For example, in alliance–outcome research, the within-therapist correlation tells us how the alliance is related to outcome within a given therapist. Thus, the within-therapist correlation tests the association between patient variability in the alliance and outcome (i.e., the therapist is a constant for patients nested within that particular therapist, ignoring the interaction for now). More simply, the within-therapist correlation tests whether patients who report a high alliance also report better outcomes than patients of the same therapist who report a low alliance. In contrast, the between-therapist correlation tells us how therapists' average alliance is related to their average outcome. That is, the between-therapist relationship tells us how variability among therapists in the alliance is related to outcome. By separating within- and between-therapist correlations, we can also test whether there is an interaction between the patients' and therapists' variability. This interaction, known as a cross-level interaction because it is the interaction between a patient-level variable (e.g., patients' rating of the alliance) and a therapist-level variable (e.g., therapists' average alliance rating; Snijders & Bosker, 1999), tells us whether the correlation between patients' variability in the alliance and outcome depends on therapists' variability in the alliance and vice versa.

Separating within- and between-therapist relationships also allows researchers to explore contextual effects (Raudenbush & Bryk, 2002). Statistically, contextual effects are the difference between the within- and between-therapist alliance–outcome correlations. Conceptually, contextual effects reveal how outcomes differ among patients who have the same alliance scores but see different therapists (Raudenbush & Bryk, 2002). Furthermore, contextual effects allow researchers to quantify how much outcomes may be improved by seeing a therapist who (on average) forms strong alliances over seeing a therapist who (on average) forms weak alliances.

Research on the differential effectiveness of therapists underscores the importance of distinguishing between within- and between-therapist relationships (Crits-Christoph et al., 1991; Crits-Christoph & Mintz, 1991; Wampold & Brown, 2005). This research typically partitions the variance in patient outcomes into two parts: (a) variability among therapists (i.e., therapist variance) and (b) variability within therapists (i.e., patient variance). Modeling the variance this way is accomplished via multilevel models (Raudenbush & Bryk, 2002; Snijders & Bosker, 1999). Crits-Christoph et al. (1991), in a seminal article, argued for the disaggregation of patient and therapist effects in psychotherapy research and provided what has turned out to be fairly robust estimates of the variability of outcomes due to therapists. Although there continues to be some disagreement about the magnitude of therapist effects in some contexts (cf. Elkin, Falconnier, Martinovich, & Mahoney, 2006; Kim, Wampold, & Bolt, 2006), multilevel models applied to clinical trial and naturalistic data have shown that 5%–10% of the variance in outcomes is due to therapists—that is, 5%–10% of the total variability in outcomes is attributable to between-therapist differences (Crits-Christoph et al., 1991; Elkin et al., 2006; Kim et al., 2006; Wampold & Brown, 2005).

In summary, in order to understand the correlation between alliance and outcome, it is critical to separate patient and therapist variability in the alliance. The primary aim of the present study was to explore the relative importance of patient and therapist variability in the alliance as they relate to outcomes. In addition, our analyses also have implications for the hypothesis that the alliance is a consequence of early symptom change. We are unaware of any alliance–outcome research that has distinguished between within- and between-therapist correlations and modeled them simultaneously. Instead, the most common type of alliance–outcome research has examined the correlation of alliance and outcomes, ignoring the fact that patients are nested within therapists—the resulting correlation is what is said to be the "total" correlation, which is in essence the average of within- and between-therapist correlations (Snijders & Bosker, 1999). Because the within- and between-therapist correlations may not be equal and conceivably could be in opposite directions, the total correlation may be misleading. Regardless, the total correlation does not provide information about the relative importance of patient and therapist variability in the alliance. Furthermore, theoreticians tend to emphasize the importance of both therapists and patients. Consequently, it is difficult to predict how the alliance is related to outcomes within and between therapists, and in that regard, the present research is exploratory.

Method

Participants and Procedures

The data analyzed in this study were patient outcomes drawn from the Research Consortium of Counseling and Psychological Services in Higher Education psychotherapy process and outcome study (see Brownson, 2004). The Research Consortium was founded in 1990 with the purpose of creating a large psychotherapy research network and presently consists of 45 university counseling centers nationwide. The Research Consortium's psychotherapy process and outcome study is a naturalistic database collected during the 1997–1998 school year. Each center involved in the study received approval of its Institutional Review Board (Brownson, 2004). Participants in the Consortium's study completed the Outcome Questionnaire-45 (OQ-45; Lambert et al., 2004)—a measure of psychotherapy outcome—at intake and prior to each session, although only patients' intake OQ-45 and final session OQ-45 were available for analysis. In addition, participants also completed the Working Alliance Inventory (WAI; Horvath & Greenberg, 1989)—a measure of the therapeutic alliance—prior to Session 4. Complete details of all measures in the Consortium's database can be found in Brownson (2004).

There were inclusion criteria related to patients and therapists. Patients had to have completed the WAI, which meant that only patients who attended at least four sessions were included. We included only therapists who saw at least two patients who had met the patient criteria.

Using these criteria, we identified 331 patients seen by 80 therapists. Patients attended an average of 7.32 sessions (SD = 3.83, range = 4–38). Therapists' average caseload was 4.1 (range = 2–18). The sample of therapists included in the Research Consortium's entire outcome study were predominately women (64%), White (79.3%), had a mean age of 37.65 (SD = 10.37), had a master's or doctoral degree (77.5%), and had, on average, 6.89 years (SD = 8.11) of experience, although demographic data on the therapists used in the sample for this study were not available. Of the 331 patients, 58.9% were women, 39.9% were men, and 8.2% did not report their gender. The patients' mean age was 23.3 (SD = 5.10, range = 17–52). The majority of patients were White (77%), followed by Hispanic (6%), Asian American (5%), African American (3%), and Native American (0.004%). Ethnicity was unknown for 9% of the patients. The sum of the percentages slightly exceeds 100% due to rounding. Diagnostic information was not available. Patients' average intake OQ-45 score was 71.27 (SD = 24.69), which is within the clinical range for the OQ-45 and comparable to other outpatient samples (Lambert et al., 2004). Furthermore, researchers have documented that counseling centers often serve patients with severe and chronic problems (Benton, Robertson, Tseng, Newton, & Benton, 2003; Erdur-Baker, Aberson, Barrow, & Draper, 2006).

Measures

Outcome measure. The primary outcome measure was the OQ-45 (Lambert et al., 2004). The OQ-45 is a self-report measure designed to assess patient outcomes during the course of therapy. The 45 items assess three primary dimensions: (a) subjective discomfort (e.g., anxiety and depression—"I feel blue"), (b) interpersonal relationships (e.g., "I feel lonely"), and (c) social role performance (e.g., "I have too many disagreements at work/school"). All 45 items may be aggregated to create a total score. Total scores can range from 0 to 180, with higher scores reflecting poorer psychological functioning. We used the OQ-45 total score in this study. The OQ-45 has been shown to have good internal consistency (α = .93), 3-week test–retest reliability (r = .84), and concurrent validity (Lambert et al., 2004; Snell, Mallinckrodt, Hill, & Lambert, 2001).

Alliance measure. The therapeutic alliance was measured with the patient-rated WAI (Horvath & Greenberg, 1989). Although the alliance is not reducible to the patient's experience of it, the patient's experience is important to understanding the relationship between alliance and outcome. Indeed, patient-rated alliance has been a consistent predictor of outcome (A. O. Horvath & Symonds, 1991). The WAI is a 36-item measure that assesses three aspects of the therapeutic relationship, proposed by Bordin (1979): (a) the bond between patient and therapist, (b) the extent to which the patient and therapist agree on the goals of treatment, and (c) the extent to which the patient and therapist see the tasks of therapy as relevant. All 36 items may be aggregated to create a total score, with high scores reflecting strong alliances. The WAI has been shown to have good internal consistency (α = .93) and adequate convergent and predictive validity (Horvath & Greenberg, 1989).

Patient and therapist variability in the alliance. The total variability in the WAI scores can be decomposed into two components (ignoring measurement error): therapist variability and patient variability. We estimated therapist variability in the alliance by calculating each therapist's mean WAI score and calculating how much each therapist's mean deviated from the WAI grand mean. We estimated patient variability in the alliance by calculating how much each patient's WAI score deviated from his or her therapist's mean WAI score. Given that we used patient-rated alliance, patient and therapist variability should be interpreted as patient and therapist variability as rated by the patient.

Statistical Analyses

We used multilevel models to address the aims of the present article. Multilevel models (Raudenbush & Bryk, 2002; Snijders & Bosker, 1999) are well suited to address the aims of our study for four reasons: (a) They take into account the hierarchical nature of therapy data, (b) they are able to consider therapists as a random factor, (c) they allow between- and within-therapist correlations to be modeled simultaneously, and (d) they use estimation procedures that are robust for unequal sample sizes within therapists. All models were estimated with Stata's (Version 9; StataCorp, 2005) xtmixed routine using restricted maximum likelihood estimation procedures.

When estimating multilevel models, an important consideration is whether to consider therapist as a fixed or random factor. Martindale (1978) noted, "The researcher who wishes to study psychotherapy is faced with the necessity of generalizing findings to two populations: a population of therapists and a population of patients" (p. 1526). Thus, when modeling therapist effects, researchers must model them in a way that allows generalizations beyond the particular therapists in the study. Such generalizations are possible when therapists are modeled as a random factor (Serlin, Wampold, & Levin, 2003; Siemer & Joorman, 2003), and, consequently, we modeled therapists as a random factor.

We estimated two sets of multilevel models: the first set with two models and the second set with three models. Models 1a and 1b deal with the relationship between pretest OQ-45 scores and posttest OQ-45 scores. Because patients were not randomly assigned to therapists but were assigned via standard case assignment procedures of the counseling centers, it was essential to statistically adjust for pretreatment differences. Model 1a was as follows:

$$Y_{ij} = \gamma_{00} + \gamma_{10}(x_{ij} - \bar{x}_j) + \gamma_{01}(\bar{x}_j - \bar{x}) + [U_{0j} + R_{ij}],$$

where Y_{ij} is the posttest OQ-45 score for the ith patient seen by the jth therapist; x_{ij} is the pretest OQ-45 score for the ith patient seen by the jth therapist; $\bar{x}_j$ is the mean pretest OQ-45 score for the jth therapist (i.e., pretest OQ-45 averaged across all of his or her patients); $\bar{x}$ is the pretest OQ-45 grand mean; γ_{00} is the intercept; γ_{10} is the within-therapist regression coefficient for pretest OQ-45; γ_{01} is the between-therapist regression coefficient for pretest OQ-45; U_{0j} is the between-therapist variance (σ^2_{ther}); and R_{ij} is the within-therapist variance (σ^2_e). Coefficients inside the brackets are the random factors, and coefficients outside the brackets are the fixed factors. As can be seen in the equation, we centered each patient's pretest OQ-45 score around his or her therapist's mean pretest OQ-45 score. Centering the pretest OQ-45 score in this way ensures that Model 1a returned both the within- (γ_{10}) and between-therapist (γ_{01}) regression coefficients (see Raudenbush & Bryk, 2002, pp. 139–141; Snijders & Bosker, 1999, Section 4.5).

Model 1b is exactly the same as Model 1a except we added a parameter that tested whether there was an interaction between patient-level pretest OQ-45 and therapist-level pretest OQ-45. This cross-level interaction is denoted γ_{11}. This interaction term was not significant and thus was dropped from all subsequent models.

Models 2a, 2b, and 2c deal with the relationship between patient-rated therapeutic alliance and posttest OQ-45 scores adjusted for pretest OQ-45. Model 2a retained the coefficients from Model 1a but also added two parameters to estimate the relationship between alliance and outcome:

$$Y_{ij} = \gamma_{00} + \gamma_{10}(x_{ij} - \bar{x}_j) + \gamma_{01}(\bar{x}_j - \bar{x}) + \gamma_{20}(z_{ij} - \bar{z}_j) + \gamma_{02}(\bar{z}_j - \bar{z}) + [U_{0j} + R_{ij}],$$

where z_{ij} is the WAI score for the ith patient seen by the jth therapist; $\bar{z}_j$ is the mean WAI score for the jth therapist (i.e., WAI averaged across all his or her patients); $\bar{z}$ is the WAI grand mean; γ_{20} is the within-therapist regression coefficient for the WAI and indexes the relationship between patient variability in the alliance and outcome; and γ_{02} is the between-therapist coefficient for the WAI and indexes the relationship between therapist variability in the alliance and outcome. Similar to Model 1a, we centered each patient's WAI score around his or her therapist's mean WAI score so that Model 2a returned both the within- (γ_{20}) and between-therapist (γ_{02}) regression coefficients for the WAI.

Model 2b is exactly the same as Model 2a except we added a parameter that tested whether there was an interaction between patient-level WAI and therapist-level WAI. Thus, the coefficient for the interaction indexes the relationship between the interaction of patient and therapist variability in the alliance and outcome. This cross-level interaction is denoted γ_{22}. This interaction term was not significant and thus was dropped from all subsequent models.

Model 2c tested whether the difference between the within- and between-therapist regression coefficients for the WAI was significant. Model 2c was the same as Model 2a except we no longer centered z_{ij} around $\bar{z}_j$, but instead centered it around $\bar{z}$, which changes γ_{02} from being the between-therapist coefficient for the WAI to being the difference between the within- and between-therapist coefficients for the WAI (Raudenbush & Bryk, 2002; Snijders & Bosker, 1999). This difference is the contextual effect for the WAI. We denote the coefficient representing the contextual effect for the WAI as $\hat{\gamma}_{02}$. As we discuss below, the contextual effect also has implications for the hypothesis that the alliance is a consequence of good outcomes.

We also estimated random slopes models in which the coefficients for patient-level variables (γ_{10} and γ_{20}) were allowed to vary across therapists. There was no evidence that slopes varied across therapists. Furthermore, the variance components for slopes were so close to zero that the models would not converge, which may be a consequence of overparameterization. Thus, we exclude random slopes from our model descriptions in Tables 1 and 2.

We used the software program PINT (Bosker, Snijders, & Guldemond, 2003; Snijders & Bosker, 1993) to calculate power to detect a within-therapist coefficient of .20 for the WAI and a between-therapist coefficient of .20 for the WAI, given our sample size of 331 patients and 80 therapists. Power exceeded 90% for the within-therapist coefficient and 80% for the between-therapist coefficient.

Results

Total Alliance–Outcome Correlation

The mean for pretest OQ-45 was 71.27 (SD = 24.69), the mean for posttest OQ-45 was 56.84 (SD = 23.33), and the mean for Session 4 patient-rated WAI was 211.63 (SD = 26.13). The correlation between the WAI and posttest OQ-45 was −.24 ($p <$ .001), indicating that higher patient-rated alliances corresponded to lower posttest OQ-45 scores. The partial correlation between the WAI and posttest OQ-45 adjusting for pretest OQ-45 was −.21 (p = .001). These correlations are similar to the average alliance–outcome correlations found in meta-analyses (Horvath & Bedi, 2002; Horvath & Symonds, 1991; Martin et al., 2000). However, as noted above, the overall alliance–outcome correlation does not separate within- and between-therapist correlations, making it difficult to understand the relative importance of patient and therapist variability in the alliance. Furthermore, within- and between-therapist correlations can be markedly different from the total relationship, as we show (cf. Snijders & Bosker, 1999).

Outcome Predicted by Pretest OQ-45

Models 1a and 1b explored the relationship between pretest OQ-45 and posttest OQ-45 within- and between therapists. As can be seen in Table 1, the results from Model 1a indicated that within therapists, there is a positive relationship between pretest OQ-45 and posttest OQ-45 (γ_{10} = 0.40, $p <$ .01). Additionally, between therapists, there was a positive relationship between pretest OQ-45 and posttest OQ-45 (γ_{01} = 0.63, $p <$ .01). The interaction between patients' pretest OQ-45 and therapists' mean pretest OQ-45 was

Table 1
Multilevel Models Predicting Posttest OQ-45 From Pretest OQ-45

Variable	Model 1a		Model 1b	
	Coefficient	95% CI	Coefficient	95% CI
Fixed effects				
Intercept (γ_{00})	56.21*	53.85, 58.56	56.89*	54.20, 59.57
Pretest OQ-45				
Within therapist (γ_{10})	0.40*	0.30, 0.51	0.41*	0.30, 0.51
Between therapist (γ_{01})	0.63*	0.47, 0.81	0.62*	0.45, 0.79
Cross-level interaction (γ_{11})	—	—	−0.01	−0.02, 0.002
Random effects				
Therapist variance (σ^2_{ther})	12.69		12.94	
Patient variance (σ^2_e)	393.53*		391.53*	
ρ_I	0.03		0.03	

Note. Patient $N = 331$; Therapist $N = 80$. OQ-45 = Outcome Questionnaire-45; CI = confidence interval; ρ_I = intraclass correlation for therapist. Dashes indicate that the cross-level interaction was not estimated in Model 1a.
* $p < .01$.

not significant ($\gamma_{11} = -0.01$, $p = .12$) and was dropped from all further analyses (see Model 1b).

Outcome Predicted by Working Alliance

Models 2a, 2b, and 2c explore the relationship between the WAI and posttest OQ-45, adjusting for pretest OQ-45. As can be seen in Table 2, in Model 2a, the relationship between pretest OQ-45 and posttest OQ-45 both within- and between therapists was similar to what was found in Model 1a.

The models exploring the relationship between outcome and patient and therapist variability in the alliance, as perceived by the patient, highlighted the importance of therapist variability. Specifically, Model 2a indicated that within therapists, there was no significant relationship between the WAI and posttest OQ-45 ($\gamma_{20} = -0.08$, $p = .11$). That is, among patients seen by the same therapist, there was no relationship between level of the working alliance and outcome, adjusting for baseline functioning. In contrast, Model 2a indicated that between therapists, there was a significant relationship between the WAI and posttest OQ-45 ($\gamma_{02} = -0.33$, $p < .01$). That is, it is predicted that patients of a therapist one standard deviation above (or below) the mean WAI for therapists (i.e., WAI averaged across a therapist's patients) would, on average, have a posttest OQ-45 score 5.8 points lower (or higher) than patients of a therapist with average WAI scores, adjusting for baseline OQ-45. Figure 1 illustrates the difference between the within- and between-therapist regression lines. More plainly, therapists whose patients, on average, rated their alliance high also had better outcomes than therapists whose patients, on average, rated their alliance low.

Model 2b indicated that the interaction between the patients' WAI score and therapists' mean WAI score was not significant

Table 2
Multilevel Models Predicting Posttest OQ-45 From Pretest OQ-45 and Patient-Rated Working Alliance

Variable	Model 2a		Model 2b		Model 2c	
	Coefficient	95% CI	Coefficient	95% CI	Coefficient	95% CI
Fixed effects						
Intercept (γ_{00})	56.60*	54.46, 58.74	56.77*	54.38, 59.15	56.60*	54.46, 58.74
Pretest OQ-45						
Within therapist (γ_{10})	0.40*	0.29, 0.50	0.40*	0.29, 0.50	0.40*	0.29, 0.50
Between therapist (γ_{01})	0.50*	0.33, 0.67	0.50*	0.33, 0.67	0.50*	0.33, 0.67
WAI (Session 4)						
Within therapist (γ_{20})	−0.08	−0.18, 0.02	−0.08	−0.18, 0.02	−0.08	−0.18, 0.02
Between therapist (γ_{02})	−0.33*	−0.49, −0.17	−0.33*	−0.50, −0.17	—	—
Cross-level interaction (γ_{22})	—	—	0.001	−0.007, 0.007	—	—
Contextual effect ($\hat{\gamma}_{02}$)	—	—	—	—	−0.25*	−0.43, −0.06
Random effects						
Therapist variance (σ^2_{ther})	0.45		0.32		0.45	
Patient variance (σ^2_e)	385.57*		386.88*		385.57*	
ρ_I	0.001		0.001		0.001	

Note. Patient $N = 331$; Therapist $N = 80$. OQ-45 = Outcome Questionnaire-45; CI = confidence interval; WAI = Working Alliance Inventory; ρ_I = intraclass correlation for therapist. Dashes indicate that the cross-level interaction was not estimated in Model 2a or 2c, that the contextual effect was not estimated in Model 2a or 2b, and that the between-therapist effect was not estimated in Model 2c.
* $p < .01$.

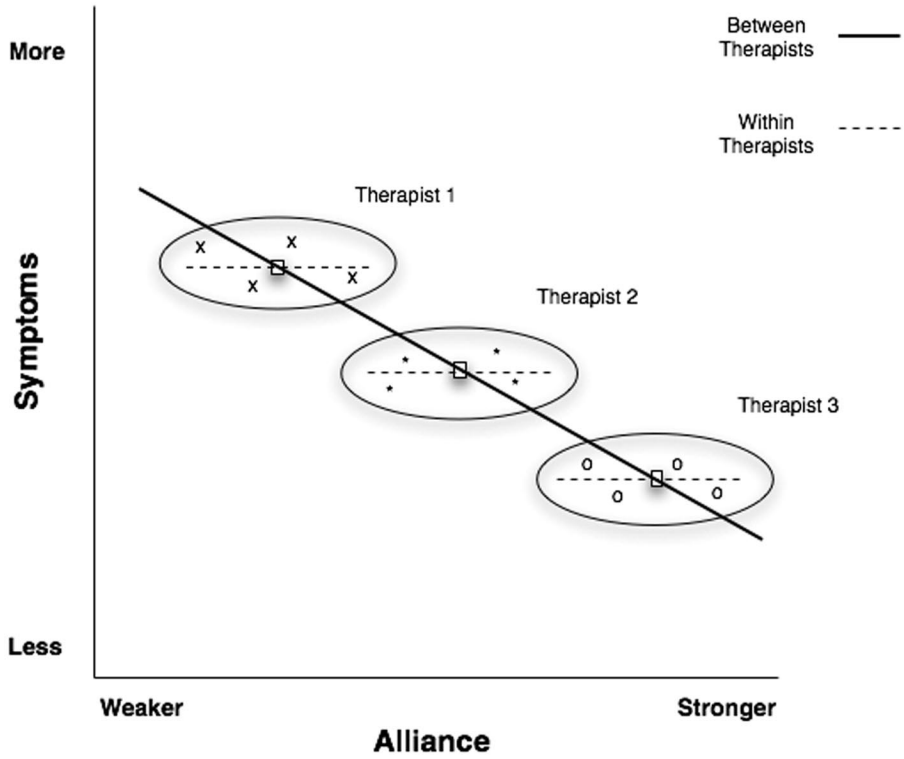

Figure 1. An illustration of the within- and between-therapist alliance–outcome correlations. Within- and between-therapist correlations are illustrated with only three therapists using simulated data to aid in the presentation of the correlations. The Xs refer to the alliance scores for Therapist 1's patients; the asterisks refer to the alliance scores of Therapist 2's patients; open circles refer to the alliance scores of Therapist 3's patients; open squares refer to each therapist's mean alliance score.

($\gamma_{22} = 0.001$, $p = .96$) and was dropped from all further analyses. Thus, there is no evidence that the alliance–outcome correlation is a consequence of an interaction between the patients' variability in the alliance and the therapists' variability in the alliance.

The results from Model 2c indicated that the contextual effect for the WAI was significant ($\hat{\gamma}_{02} = -0.25$, $p < .01$), and thus the between- and within-therapist coefficients were statistically different from each other. The contextual effect can be interpreted as the expected difference in posttest OQ-45 (adjusted for pretest OQ-45) between patients with the same WAI score but who see therapists who differ by one unit in their mean WAI score. Thus, for two patients who have the same WAI score but who see therapists who differ by one standard deviation in their mean WAI score, the difference in their posttest OQ-45 adjusted for pretest OQ-45 is 4.4, which is 22.2% of a standard deviation of the adjusted posttest OQ-45. Furthermore, the difference in adjusted posttest OQ-45 between a patient seen by a therapist one standard deviation below the WAI mean and a patient seen by a therapist one standard deviation above the mean is 8.8, which is 44.4% of a standard deviation of the adjusted posttest OQ-45. Figure 2 provides an illustration of the contextual effect for patients and therapists with average pretest OQ-45 scores.

Alliance-as-Outcome Hypothesis

The contextual analysis also has implications for DeRubeis et al.'s (2005) fourth explanation of the alliance–outcome correlation. DeRubeis et al.'s (2005) fourth explanation, which we call the "alliance-as-outcome" hypothesis, is that the alliance is partially or even wholly an effect of previous symptom reduction rather than the cause of symptom reduction (see also Crits-Christoph, Connolly Gibbons, & Hearon, 2006). Thus, alliance is at least in part a pseudooutcome, and the alliance–outcome correlation represents the correlation between two outcome measures. The alliance-as-outcome hypothesis makes a specific prediction about the relative magnitude of the within- and between-therapist alliance–outcome regression coefficients. If the alliance is a function of previous symptom change, then the patient and therapist variability in the alliance would not be differentially related to outcome, and thus the within- and between-therapist coefficients would be approximately equal. Within a therapist, those patients who have experienced the most initial symptom change should experience relatively positive alliance and final outcome (i.e., a significant within-therapist alliance and final outcome correlation). As well, the most skilled therapists should demonstrate more initial change with their patients generally, which would lead to stronger alliance ratings and better final outcomes (i.e., a significant between-therapist alliance and final outcome correlation). Therefore, what is most important is how much change has occurred up to the point the alliance was measured, regardless of whether the patient is nested within a particular therapist. Consequently, the within- and between-therapist coefficients

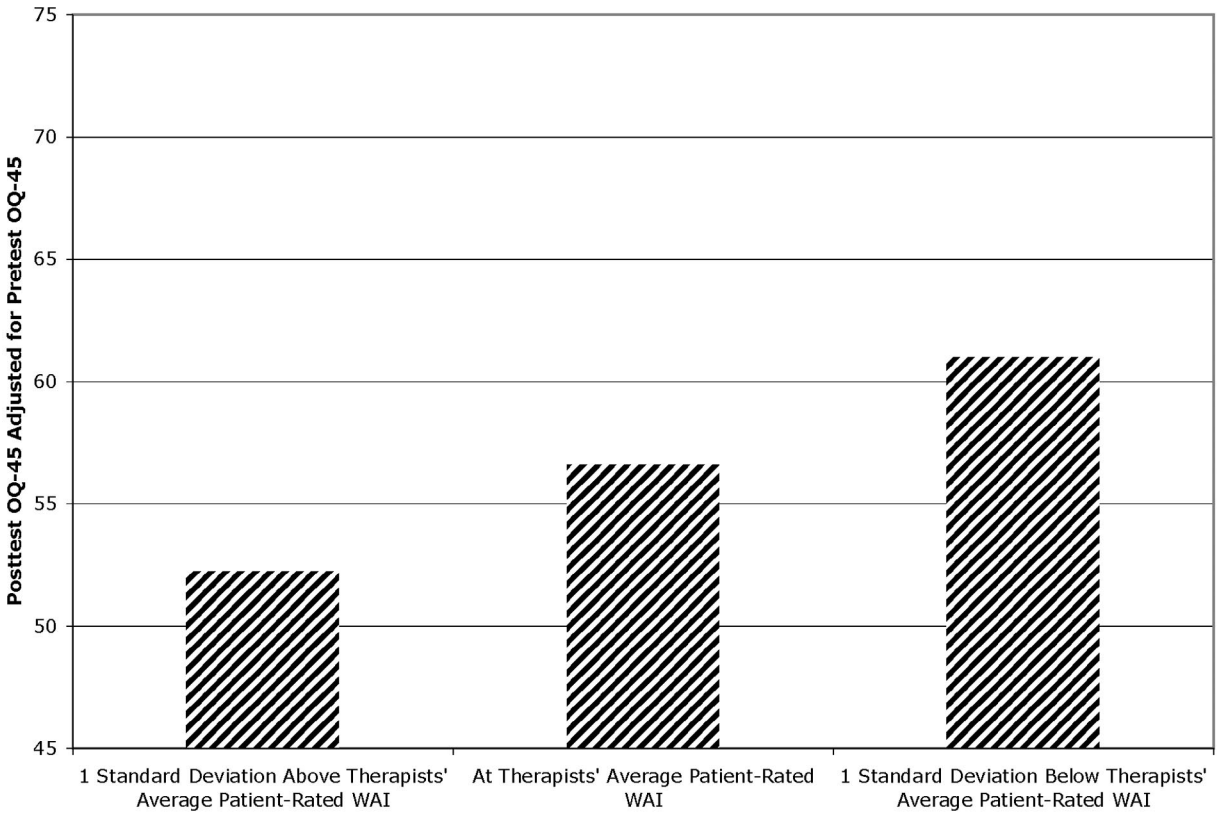

Figure 2. Predicted posttest Outcome Questionnaire-45 (OQ-45) scores as a function of therapists' average patient-rated Working Alliance Inventory (WAI) score and pretest OQ-45. Predicted scores were generated for therapists whose mean WAI score (i.e., averaged across all their patients) was at the mean of all therapists, one standard deviation below the mean of all therapists, and one standard deviation above the mean of all therapists. Predicted scores were generated at the patient mean and therapist mean for pretest OQ-45 and the patient mean for WAI.

would be approximately equivalent to each other and to the overall alliance–outcome correlation.

The contextual analysis (Model 2c) does not support the alliance-as-outcome hypothesis. The contextual analysis provides a test of whether the within- and between-therapist regression coefficients are significantly different from one another. The results of the contextual analysis suggested that the within- and between-therapist regression coefficients differed significantly. Within therapists, final outcome was not related to the alliance. If alliance is a consequence of initial symptom change, even partially, then those patients with higher alliances should have better outcomes. In contrast, we observed that within therapists those patients with relatively high alliance ratings did not have better outcomes than those patients with relatively low alliance ratings. Furthermore, between therapists, final outcome was related to the alliance. Those therapists who generally had high alliance ratings also had better outcomes than those therapists with generally low alliance scores. The contextual effect also tests whether two people who have the exact same total alliance score but who see therapists with different average alliance scores will have different outcomes. If alliance is measuring symptom reduction, then patients with the same alliance score, which is due systematically to previous symptom reduction, should have the same predicted outcome, all other variables held constant. However, the results of the contextual analysis suggested that even when two patients have the exact same total alliance score, they are predicted to have different outcomes if they see therapists with different average alliance scores. Therefore, the results of the contextual analysis suggest that the alliance–outcome correlation is more than an artifact of prior symptom change. Indeed, the contextual analysis suggests that the therapist variability in the alliance, apart from the initial change in symptoms, is important to outcome.

Accounting for Therapist Effects

Multilevel models may also be used to explore whether the alliance mediates the relationship between therapists and outcome by determining whether therapist effects are reduced when alliance is entered in the model. We used the intraclass correlation for therapist (ρ_I) as an index of therapist effects. ρ_I is calculated as follows:

$$\rho_I = \frac{\sigma^2_{ther}}{\sigma^2_{ther} + \sigma^2_e},$$

where σ^2_{ther} is the between-therapist variance and σ^2_e is the within-therapist variance (i.e., patient variance). Thus, the intraclass correlation can be interpreted as the proportion of variance in the outcome due to therapists.

Using the information from Model 1a, we calculated ρ_I as follows:

$$\frac{\sigma^2_{ther}}{\sigma^2_{ther} + \sigma^2_e} = \frac{12.69}{12.69 + 393.53} = 0.03.$$

Thus, after adjusting for pretest OQ-45, 3% of the variance in posttest OQ-45 was due to therapists. However, ρ_I from Model 1a is not statistically significant. Consequently, the following analysis should be considered illustrative and interpreted accordingly. Future researchers should consider using similar methods to explore whether relevant variables (e.g., alliance, treatment adherence) account for significant therapist effects.

Using the information from Model 2a, we again calculated ρ_I as follows:

$$\frac{\sigma^2_{ther}}{\sigma^2_{ther} + \sigma^2_e} = \frac{0.45}{0.45 + 385.57} = 0.001.$$

Thus, adjusting for pretest OQ-45 and, importantly, the WAI, therapists accounted for 0.1% of the variance in posttest OQ-45. Thus, differences among therapists in patient-rated alliance reduced the therapist effects by 97% (i.e., from 3% to 0.1%; Singer, 1998).

Discussion

The ubiquitous relationship between alliance and outcomes is typically reported as a total correlation, which fails to separate therapist and patient variability in the alliance. In the present study, we modeled the relationship of alliance and outcome both within- and between therapists to separate therapist and patient variability in the alliance. We found that therapist and patient variability in patient-rated alliance do not equally predict patient outcomes. Specifically, the results of the study indicated that therapist variability in the patient-rated alliance accounted for the alliance–outcome correlation. That is, therapists who, on average, formed stronger alliances with their patients showed statistically significant better outcomes than therapists who did not form as strong of alliances (see Figures 1 and 2). Furthermore, within the caseload of a given therapist, the strength of the alliance did not significantly predict outcome, which suggests that patient variability in the alliance may be unimportant to outcome (see Figure 1). There was no evidence of a Patient × Therapist interaction. Finally, the significant difference between the within- and between-therapist regression coefficients suggested that the alliance is not simply a consequence of early symptom change.

If, as the results of this research suggest, it is the therapist variability in the alliance that is related to outcome, then what characteristics or actions of the therapist are responsible for the alliance and the subsequent outcomes? In a review of the alliance literature, Ackerman and Hilsenroth (2003) identified a number of therapist attributes and therapist techniques that have been correlated with strong alliances, including being flexible, experienced, honest, respectful, trustworthy, confident, interested, alert, friendly, warm, and open. Important therapist techniques included being reflective, being supportive, noting past therapy success, providing accurate interpretations, facilitating emotional expression, being active, and being affirming (see Ackerman & Hilsenroth, 2003, Table 3, p. 28). Each of these attributes and techniques could vary between therapists and thus account for the observed between-therapist differences in the alliance. Future theoretical and empirical work in this area should aim to provide therapists with clear definitions of each attribute and technique and how therapists can develop the attributes and use the techniques. An excellent example of this type of theoretical work is Linehan's (1997) writings on validation and its effects on the therapeutic relationship. Linehan provided detailed descriptions of six forms (levels) of validation and identified some of the contexts in which a therapist would use each of the forms of validation.

Therapist variability in the alliance may be related to therapists' ability to manage and repair ruptures in the alliance (Safran & Muran, 1996, 2000). *Ruptures* refers to reductions in the alliance over time. Some have described the pattern of the alliance over time as one in which there is a rupture in the alliance followed by a repair to the alliance (Safran & Muran, 2000). Safran and Muran (2000) have developed a treatment guide aimed at specifying productive strategies for managing and repairing ruptures in the therapeutic alliance (see also Burns & Auerbach, 1996, for a related approach for dealing with ruptures in the alliance). Future research could explore whether therapists' ability to competently and efficiently repair ruptures in the therapeutic alliance determines their ability to form and maintain strong relationships with their patients.

The finding that therapist variability drives the alliance–outcome correlation is also consistent with common factors theory. For example, effective therapists are thought to build and augment the alliance by fostering hope and providing a treatment that is coherent and responsive to patients' needs (Frank & Frank, 1991). Effective therapists may foster an expectation among their patients that treatment can and will help by providing their patients coherent and relevant explanations of their patients' problems and delivering a treatment consistent with such explanations. This may help patients feel understood and collaborate on the tasks and goals of therapy, thus enhancing the probability of success and building the alliance. In turn, a strong alliance will increase the likelihood that patients will agree with their therapists on tasks and goals of therapy, which will affect outcomes (Wampold, Imel, Bhati, & Johnson, 2006). Therapists will likely vary with regard to their skill in instilling hope, providing coherent explanations of patients' problems, and delivering treatments consistent with their explanations. Therefore, there will be differences between therapists in their average alliance scores, which, as our results showed, are associated with patient outcomes.

Many theorists have conjectured that patients are largely responsible for the alliance, and variability among patients in the alliance is related to outcome (e.g., Mallinckrodt, 2000). If this were the case, then within-therapist alliance–outcome correlations would be expected, a hypothesis not corroborated in the present study. We propose two explanations for why we do not see a within-therapist alliance–outcome relationship.

The first explanation is methodological, although it has theoretical implications. The alliance was measured cross-sectionally and at an arbitrary point (Session 4). If the longitudinal trajectory of the alliance is one of rupture and repair (Horvath & Marx, 1991;

Safran & Muran, 2000), then we would expect some ruptures to be occurring at Session 4 when the alliance was measured. For other cases in which the patient had not experienced a rupture or in which the therapist and patient had repaired the rupture, the alliance would be relatively high. Thus, the arbitrary chosen time point results in an alliance measurement that may not be indicative of the general nature of the alliance for that case. The cross-sectional nature of the measurement of the alliance and variability over time of the alliance would tend to attenuate within-therapist correlations. However, it should be noted that the cross-sectional design does not invalidate the between-therapist relationship because therapists who generally repair their alliances will exhibit higher alliances cross-sectionally than therapists who do not repair their alliances (recall that the therapist alliance score is the average of the alliance ratings of all the patients a particular therapist saw).

A second explanation for the lack of a within-therapist correlation can be understood from a rater bias perspective. Patient variability in the alliance rating may in part represent raters' tendencies to be complimentary or critical of relationships with their therapists, regardless of what the therapist contributes to the relationship (Hoyt, 2002). For example, consider two therapists—Therapist A and Therapist B. Therapist A, on average, has high alliance ratings and Therapist B, on average, has low alliance ratings. Also consider Patient X, who comes to therapy with a tendency to be critical of relationships. We would expect Patient X's alliance ratings of Therapist A to be higher than Therapist B because Therapist A is better able to form alliances. However, because Patient X is generally critical of relationships, Patient X's ratings of Therapist A would be lower than the rest of Therapist A's patients. The same would be true of Patient X's ratings of Therapist B. Patient X would benefit (or be hindered) from his or her therapist's ability to form an alliance generally, although Patient X may perceive the alliance as slightly worse than it really is because of his or her tendency to be critical. In any event, Patient X's relative low rating of either therapist does not seem to be related to how much Patient X will benefit from therapy. Likewise, Patient Y, with the opposite tendency (i.e., generally would give high ratings), would not be expected to have a better outcome than Patient X if they both see the same therapist. Therefore, some of the within-therapist variation may be due to patients' rating tendencies, and the extent to which these tendencies are unrelated to outcome is the extent to which within-therapist variation is unrelated to outcomes (see Hoyt, 2002, for a discussion of how to estimate the magnitude of this rater bias).

We did not observe a relationship between outcome and the interaction between patient variability and therapist variability in the alliance. Thus, it does not appear that the between-therapist alliance–outcome correlation is moderated by their patients' variability in the alliance. Of course, it is possible that therapists vary in their success in forming alliances, depending on the type of patient. However, this interaction between patient and therapist may be unrelated to outcome, as our results suggested. We also did not observe any evidence that the within-therapist alliance–outcome relationship varied across therapists (i.e., no evidence of a random slope). This indicated that the observed null relationship between alliance and outcome adjusted for pretest symptoms is constant across therapists. That is, it is not the case that there is a significant within-therapist alliance–outcome correlation for some therapists but not others.

Our results were not consistent with the contention that alliance is a consequence of early change, for two reasons. First, we did not observe a within-therapist alliance–outcome correlation, which would be expected if alliance was a function of early reduction in symptoms, as those patients with the most early change in a therapist's caseload would also produce the highest alliance ratings and the most overall change. Second, the results of the contextual analysis (Model 2c) suggested that two patients with the exact same alliance rating would have different outcomes depending on what therapist they worked with, which does not support the idea that what is most important is how much change has occurred up to the point the alliance was measured, regardless of whether the patient is nested within a particular therapist. Therefore, the results of this study suggest, albeit indirectly, that the alliance–outcome correlation is not simply an artifact of early symptom change.

We were not able to address directly the alliance-as-outcome hypothesis by examining early symptom change. It should be noted that such analyses have produced mixed results (Barber, Connolly, Crits-Christoph, Gladis, & Siqueland, 2000; DeRubeis & Feely, 1990; Feeley, DeRubeis, & Gelfand, 1999; Klein et al., 2003; Tang & DeRubeis, 1999), which might be due to the fact that these investigations have not separated within- and between-therapist relationships. In addition, controlling for early symptom change is not a straightforward issue (cf. Meehl, 1971). If it is true that the alliance is necessary for change and alliance from early sessions (e.g., Sessions 1, 2, and 3) is responsible for the changes typical of early sessions, then adjusting for early symptom change may partial out important variance in the alliance, making it impossible to find a relationship between alliance and outcome. Thus, the statistical controls would, as Meehl (1971) said, "pseudofalsify a good causal theory" (p. 143). Future research on the alliance–outcome correlation would benefit from considering multiple methods for testing the alliance-as-outcome hypothesis.

A limitation of the present study is that patients were not randomly assigned to therapists. We adjusted for pretreatment differences on the outcome measure at the patient and therapist level. Thus, the models adjust for the fact that therapists' caseloads differed with regard to pretreatment symptoms severity. Nevertheless, the influence of selection biases due to other unmeasured variables is unknown. A second limitation is that the study included only one measure of outcome and one measure of alliance. Consequently, it is unclear whether these results will generalize to other measures of outcome and alliance. Furthermore, patients rated each measure. Within- and between-therapist alliance–outcome correlations may vary by who rates the alliance and outcome. Future research should measure the alliance and outcome from multiple perspectives. A third limitation is that the sample of participants was predominately White and was drawn from university counseling centers. It is unclear whether these results would generalize to more ethnically diverse samples and other clinical populations. A fourth limitation is that we do not have complete descriptive information for patients and therapists. Consequently, we are not able to explore whether the results are moderated by patient and therapist characteristics, such as diagnosis and therapeutic orientation, respectively. Future research in this area should make efforts to collect such information. Finally, a fifth limitation is that we do not have information regarding the specific treatment delivered to each patient, a common limitation of naturalistic data. Thus, we were not able to explore whether the

between- and within-therapist correlations differed across treatment type (cf. DeRubeis et al., 2005).

Implications

The results of this study have potentially important clinical implications. Foremost among them is the relative importance of therapist variability in the alliance with regard to outcomes. In situations in which therapists have trouble forming an alliance, it would behoove therapists to attend to their own contributions to the alliance and focus less on characteristics of the patient that impede the development of the alliance. Indeed, therapist attributions of resistance or maladaptive attachment styles as an explanation of a poor alliance, according to our findings, would be irrelevant with regard to outcomes, although these explanations may be grist for therapeutic work. Along these lines, therapists may benefit from regularly monitoring their alliances and, when their alliances are substantially low, reflect on their actions and seek remedies. Monitoring alliances would require benchmarks for good alliance ratings against which any given alliance rating could be compared to determine whether a particular alliance is "on course." Similar systems for monitoring patient outcomes and comparing them with norms have been used successfully to provide therapists feedback regarding occasions when their patients are not changing at an expected rate (e.g., Lambert et al., 2003).

When therapists find that their alliance ratings are consistently low, they may benefit from supervision or further training. Crits-Christoph, Connolly Gibbons, Crits-Christoph, et al. (2006) have developed a training program aimed at helping therapists improve their alliances. Although their results are suggestive but not statistically significant, Crits-Christoph et al. found that alliance training produces moderate to large increases in alliances. Safran and Muran (2000) have argued that therapists may benefit from further training in monitoring and managing ruptures in the alliance and are presently investigating the efficacy of such training.

The results of the study also have important research implications. Specifically, the results underscore the importance of considering within- and between-therapist correlations because those correlations—and the processes that drive them—may differ (cf. Snijders & Bosker, 1999). Consequently, the total correlations—which in essence are a crude average of between- and within correlations—may provide misleading results. Multilevel models provide researchers with an excellent statistical tool for separating within- and between-therapist correlations. We recommend that researchers use multilevel models or other innovative methods whenever they are studying variables that could conceivably vary among therapists, such as the alliance, treatment adherence, and treatment competence. Doing so will provide a clearer picture of the ways in which therapists and patients contribute to the change process.

References

Ackerman, S. J., & Hilsenroth, M. J. (2003). A review of therapist characteristics and techniques positively impacting the therapeutic alliance. *Clinical Psychology Review, 23,* 1–33.

Barber, J. P., Connolly, M. B., Crits-Christoph, P., Gladis, L., & Siqueland, L. (2000). Alliance predicts patients' outcome beyond in-treatment change in symptoms. *Journal of Consulting and Clinical Psychology, 68,* 1027–1032.

Benton, S. A., Robertson, J. M., Tseng, W.-C., Newton, F. B., & Benton, S. L. (2003). Changes in counseling center client problems across 13 years. *Professional Psychology: Research and Practice, 34,* 66–72.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research, and Practice, 16,* 252–260.

Bosker, R. J., Snijders, T. A. B., & Guldemond, H. (2003). *User's manual for PINT-Power in two level designs* (Version 2.1). Gronigen, the Netherlands: Author.

Brownson, C. (2004). *The Research Consortium of Counseling and Psychological Services in Higher Education: Report of activities*. Retrieved May 11, 2007, from http://www.utexas.edu/student/cmhc/research/rescon.html

Burns, D. D., & Auerbach, A. (1996). Therapeutic empathy in cognitive-behavioral therapy: Does it really make a difference? In P. M. Salkovskis (Ed.), *Frontiers of cognitive therapy* (pp. 135–164). New York: Guilford Press.

Castonguay, L. G., & Holtforth, M. G. (2005). Change in psychotherapy: A plea for no more "nonspecific" and false dichotomies. *Clinical Psychology: Science and Practice, 12,* 198–201.

Craighead, W. E., Bjornsson, A. S., & Amarson, E. O. (2005). Specificity and nonspecifity in psychotherapy. *Clinical Psychology: Science and Practice, 12,* 189–193.

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

Crits-Christoph, P., Connolly Gibbons, M. B., Crits-Christoph, K., Narducci, J., Schamberger, M., & Gallop, R. (2006). Can therapists be trained to improve their alliances? A preliminary study of alliance-fostering psychotherapy. *Psychotherapy Research, 16,* 268–281.

Crits-Christoph, P., Connolly Gibbons, M. B., & Hearon, B. (2006). Does the alliance cause good outcome? Recommendations for future research on the alliance. *Psychotherapy: Theory, Research, Practice, Training, 43,* 280–285.

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

DeRubeis, R. J., Brotman, M. A., & Gibbons, C. J. (2005). A conceptual and methodological analysis of the nonspecifics argument. *Clinical Psychology: Science and Practice, 12,* 174–183.

DeRubeis, R. J., & Feely, M. (1990). Determinants of change in cognitive therapy for depression. *Cognitive Therapy and Research, 14,* 469–482.

Elkin, I., Falconnier, L., Martinovich, Z., & Mahoney, C. (2006). Therapist effects in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Psychotherapy Research, 16,* 144–160.

Erdur-Baker, O., Aberson, C. L., Barrow, J. C., & Draper, M. R. (2006). Nature and severity of college students' psychological concerns: A comparison of clinical and nonclinical national samples. *Professional Psychology: Research and Practice, 37,* 317–323.

Feeley, M., DeRubeis, R. J., & Gelfand, L. A. (1999). The temporal relation of adherence and alliance to symptom change in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 67,* 578–582.

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

Gelso, C. J., & Carter, J. A. (1994). Components of the psychotherapy relationship: Their interaction and unfolding during treatment. *Journal of Counseling Psychology, 41,* 296–306.

Hatcher, R. L., & Barends, A. W. (2006). How a return to theory could help alliance research. *Psychotherapy: Theory, Research, Practice, Training, 43,* 292–299.

Horvath, A. O., & Bedi, R. P. (2002). The alliance. In J. C. Norcross (Ed.),

Psychotherapy relationships that work: Therapist contributions and responsiveness to patients (pp. 37–70). New York: Oxford University Press.

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology, 36,* 223–233.

Horvath, A. O., & Marx, R. W. (1991). The development and decay of the working alliance during time-limited counseling. *Canadian Journal of Counseling, 24,* 240–259.

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

Hoyt, W. T. (2002). Bias in participant ratings of psychotherapy process: An initial generalizability study. *Journal of Counseling Psychology, 49,* 35–46.

Kazdin, A. E. (2005). Treatment outcomes, common factors, and continued neglect of mechanisms of change. *Clinical Psychology: Science and Practice, 12,* 184–188.

Kim, D.-M., Wampold, B. E., & Bolt, D. M. (2006). Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychotherapy Research, 16,* 161–172.

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

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

Lambert, M. J., Morton, J. J., Hatfield, D., Harmon, C., Hamilton, S., Reid, R. C., et al. (2004). *Administration and scoring manual for the OQ-45.2*. Orem, UT: American Professional Credentialing Services.

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

Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice, 10,* 288–301.

Linehan, M. M. (1997). Validation and psychotherapy. In A. C. Bohart & L. S. Greenberg (Eds.), *Empathy reconsidered: New directions* (pp. 353–392). Washington, DC: American Psychological Association.

Mallinckrodt, B. (2000). Attachment, social competencies, social support, and interpersonal process in psychotherapy. *Psychotherapy Research, 10,* 239–266.

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

Martindale, C. (1978). The therapist-as-fixed-effect fallacy in psychotherapy research. *Journal of Consulting and Clinical Psychology, 46,* 1526–1530.

Meehl, P. E. (1971). High school yearbooks: A reply to Schwarz. *Journal of Abnormal Psychology, 77,* 143–148.

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

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology, 21,* 95–103.

Safran, J. D., & Muran, J. C. (1996). The resolution of ruptures in the therapeutic alliance. *Journal of Consulting and Clinical Psychology, 64,* 447–458.

Safran, J. D., & Muran, J. C. (2000). *Negotiating the therapeutic alliance: A relational treatment guide*. New York: Guilford Press.

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

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

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 24,* 323–355.

Snell, M. N., Mallinckrodt, B., Hill, R. D., & Lambert, M. J. (2001). Predicting counseling center clients' response to counseling: A 1-year follow-up. *Journal of Counseling Psychology, 48,* 463–473.

Snijders, T., & Bosker, R. (1993). Standard errors and sample sizes in two-level research. *Journal of Educational Statistics, 18,* 237–260.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and multilevel analysis*. London: Sage.

StataCorp. (2005). Stata statistical software: Release 9. College Station, TX: Author.

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology, 67,* 894–904.

Wampold, B. E. (2005). Establishing specificity in psychotherapy scientifically: Design and evidence issues. *Clinical Psychology: Science and Practice, 12,* 194–197.

Wampold, B. E., & Brown, G. S. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73,* 914–923.

Wampold, B. E., Imel, Z. E., Bhati, K. S., & Johnson, M. D. (2006). Insight as a common factor. In L. G. Castonguay & C. E. Hill (Eds.), *Insight in psychotherapy* (pp. 119–140). Washington, DC: American Psychological Association.

Received November 29, 2006
Revision received May 23, 2007
Accepted May 24, 2007 ■