Journal of Consulting and Clinical Psychology
2010, Vol. 78, No. 5, 635–645

0022-006X/10/$12.00 DOI: 10.1037/a0020051

The Alliance in Couple Therapy: Partner Influence, Early Change, and Alliance Patterns in a Naturalistic Sample

Morten G. Anker
Familievernkontoret i Vestfold

Jesse Owen
University of Louisville

Barry L. Duncan
Heart and Soul of Change Project

Jacqueline A. Sparks
University of Rhode Island

Objective: The purpose of this study was to explore the relationship between the alliance and outcome in couple therapy and examine whether the alliance predicted outcomes over and above early change. The authors also investigated partner influence and gender and sought to identify couple alliance patterns that predicted couple outcomes. ***Method:*** The authors examined the alliances and outcomes at posttreatment and follow-up of 250 couples seeking treatment for marital distress in a naturalistic setting. The Session Rating Scale was used to measure the alliance; the Outcome Rating Scale and Locke Wallace Marital Adjustment Scale were used to measure outcomes. Couples were White, Euro-Scandinavian, and heterosexual, with a mean age of 38.5 years and average number of years together of 11.8. On a subsample (n = 118) that included couples with 4 or more sessions, the authors investigated the relationship between the alliance and outcome controlling for early change, and patterns of alliance development were delineated. ***Results:*** In the full sample, first-session alliances were not predictive of outcomes, but last-session alliances were predictive for both individuals and their partners. In the subsample, third-session alliances predicted outcome significantly above early change (d = 0.25) that exceeded the reliable change index. Couple alliances that started over the mean and increased were associated with significantly more couples achieving reliable or clinically significant change. Gender influences were mixed. ***Conclusions:*** Given the current findings suggesting a potential alliance impact over and above symptom relief as well as the importance of ascending alliance scores, continuous assessment of the alliance appears warranted.

Keywords: couple therapy, early change–alliance-relationship, actor–partner interdependence model, alliance patterns

Depending on the meta-analysis, the amount of outcome variance attributed to the alliance ranges from 2.5% (Horvath & Bedi, 2002), to 5% (Martin, Garske, & Davis, 2000), to 7% (Horvath & Symonds, 1991). The relationship between the alliance and outcome is remarkably consistent across treatment modalities and clinical presentations (Castonguay & Beutler, 2005). Couple therapy is no exception. For example, in a study of emotionally focused couple therapy, the alliance with the therapist explained up to 22% of outcome variance at posttreatment and 29% at follow-up (Johnson & Talitman, 1997). An investigation of 40 couples completing at least eight sessions of integrative problem-centered therapy found that the alliance accounted for 5% of the variance for men and 17% of the variance for women in marital distress at Session 8 (Knobloch-Fedders, Pinsof, & Mann, 2007). Strong alliances have been associated with greater retention (Knobloch-Fedders, Pinsof, & Mann, 2004; Raytek, McGrady, Epstein, & Hirsch, 1999), and improvement in marital distress for couples undergoing group marital skills training (Bourgeois, Sabourin, & Wright, 1990), group treatment for partner abuse (Brown & O'Leary, 2000), and behavioral marital therapy (Holtzworth-Munroe, Jacobson, DeKlynen, & Whisman, 1989).

Gender weighs in as an important consideration in couple therapy alliances although the findings are mixed. Gender has typically been addressed via separate analyses for men and women. For instance, Bourgeois et al. (1990) found that men's alliance ratings at the third session more strongly predicted outcome than women's. Knobloch-Fedders et al. (2007) reported that when men scored the alliance higher than their partners at mid-treatment (Session 8), couples showed greater improvement in marital distress. Similarly, Symonds and Horvath (2004) found that the correlation between alliance and outcome was greater for couples when men rated the alliance higher than women did (Symonds & Horvath, 2004).

Morten G. Anker, Bufetat, Familievernkontoret i Vestfold, Tønsberg, Norway; Jesse Owen, Department of Educational and Counseling Psychology, University of Louisville; Barry L. Duncan, Heart and Soul of Change Project, Tamarac, Florida; Jacqueline A. Sparks, Department of Human Development and Family Studies, University of Rhode Island.

Barry L. Duncan is the co-holder of the copyrights to the Outcome Rating Scale and the Session Rating Scale.

Correspondence concerning this article should be addressed to Barry L. Duncan, Heart and Soul of Change Project, 8611 Banyan Ct., Tamarac, Florida 33321. E-mail: barrylduncan@comcast.net

On the other hand, women's perceptions of their partner's or the couple's alliance with the therapist have shown greater predictive ability than men's perceptions. In Knobloch-Fedders et al. (2007), women's ratings of the couple's alliance at mid-treatment uniquely predicted improvement beyond that accounted for by early alliance ratings. In addition, Pinsof, Zinbarg, and Knobloch-Fedders (2008) found that women's first-session alliance scores predicted eighth-session individual and couple outcomes; these authors also reported that women's first-session perceptions of their partners' and couples' alliances predicted change in ratings of sexual dissatisfaction. Finally, Knobloch-Fedders et al. (2007) found that both men's and women's early and mid-treatment alliance scores predicted improvement. In sum, gender-linked associations between the alliance and outcome in couple therapy have surfaced for men and women. A better understanding might emerge if both partners' views of the alliance were considered simultaneously and partner influence was examined directly.

Alliance development over time is an area that has received limited study, especially in couple therapy. Symonds and Horvath (2004) compared outcomes for alliance improvers (alliance scores increased from first to third sessions) with alliance deteriorators (alliance scores decreased from first to third sessions) by gender. When both members in a couple improved, alliance correlated with outcome for men. There were no significant correlations with outcome when one partner improved and the other deteriorated or when both partners deteriorated.

Examining changes in the alliance at two points, while meaningful, may not fully capture important alliance dynamics. Some preliminary evidence from individual psychotherapy investigations suggests that different patterns of alliance development (e.g., linear, quadratic, brief V-shaped) may be associated with positive outcome (Kivlighan & Shaughnessy, 2000; Patton, Kivlighan, & Multon, 1997; Stiles et al., 2004). No patterns have been identified in couple therapy beyond what has been termed *split alliances*, in which one member of a couple is determined to differ substantially in alliance ratings with the therapist from his or her partner (see Pinsof & Catherall, 1986; Symonds & Horvath, 2004). Split alliances are typically defined at a particular session and subsequently do not reflect changes over time or patterns in the alliance.

Despite the mosaic of findings, the evidence is convincing that the alliance with the therapist is related to the success or failure of couple treatment. Considerable controversy, however, exists about the nature of the relationship and whether the association of the alliance to outcome can be separated from early symptom change. For example, some investigators examining individual therapy have found that the alliance predicts outcome when controlling for previous change (e.g., Barber, Connolly Gibbons, Crites-Christoph, Gladis, & Siqueland, 2000; Klein et al., 2003), suggesting that the alliance–outcome association does not arise solely from client improvement, while other researchers have found that the alliance has failed to correlate significantly with outcome when accounting for prior symptom change (e.g., DeRubeis & Feeley, 1990; Feeley, DeRubeis, & Gelfand, 1999). The sparse and inconsistent findings led Barber (2009) to suggest that there is not much current support that the alliance is related to further improvement in symptoms. In addition, other factors (e.g., client characteristics) are likely involved and have not been adequately studied. Highlighting the point, we know of no studies in which early change or other intervening variables and the alliance-outcome relationship have been examined in couple therapy.

In the current study, we explored the alliance in couple therapy and the impact of partner influence and gender on outcome at posttreatment and follow-up of 250 couples treated for marital distress in a routine clinical setting. In addition, on a subsample of couples attending four or more sessions ($n = 118$), we examined the relationship between early change, the alliance, and outcome. To further address couple outcomes, we also identified couple alliance patterns and their association with whether each person in a couple met reliable or clinically significant change criteria.

Specifically, we predicted that individual and partner alliance with the therapist after the first (Hypothesis 1a) and last (Hypothesis 1b) sessions would be positively related to therapy outcomes at posttreatment and follow-up. In other words, an individual's alliance with the therapist after the first and last sessions may predict not only his or her own outcomes but also the partner's outcomes. Given the mixed evidence regarding how gender may moderate the relationship between alliance and outcome, we did not make a formal prediction. Next, using a subsample that allowed for an examination of both early change and alliance patterns, we hypothesized that individual and partner alliance with the therapist at Session 3 would predict outcome over and above early treatment change (Hypothesis 2). Finally, we predicted that different patterns of couple alliance development would differentiate couple outcomes: whether none, one, or both partners achieved reliable or clinically significant change (Hypothesis 3).

Method

Participants

A total of 918 individuals or 459 couples seeking outpatient couple therapy services at two family counseling agencies providing free government-subsidized services in Norway from October 2005 to December 2007 constituted the sample pool. The final sample of 250 couples included 99 couples from the experimental feedback condition from a randomized clinical trial (RCT) in which a continuous feedback intervention was compared with treatment as usual (TAU; Anker, Duncan, & Sparks, 2009). Couples in the TAU condition (118 couples) from the RCT were eliminated because they did not complete alliance measures. The other 151 couples came from the same pool but from a second family counseling agency not participating in the RCT. Couples were excluded at phone intake if one member refused to attend, one or both members of the couple expressed the desire to end the relationship, or one or both refused informed consent. Couples were required to have attended at least two conjoint sessions and have completed the outcome and alliance measures for a minimum of the first and last sessions. This eliminated 70 couples. If the couple did not attend the last session together, the session closest to the last session that both attended was used (30 couples). Of the remaining 542 clients, 42 either did not attend the first session together (11 couples) or had no alliance and outcome scores from the same session at posttreatment (10 couples). The final sample consisted of 500 clients (250 couples).

Couples were White, Euro-Scandinavian, and heterosexual who were on average 38.54 years old ($SD = 8.47$; range from 22 to 72). Three hundred and fifty-two (71.5%) participants were employed

full time and 53 (10.8%) were employed part time, whereas 87 (17.7%) were unemployed or did not work outside the home. Eight individuals did not provide this information. Regarding education levels, 142 (28.9%) completed lower secondary school, 167 (33.9%) completed upper secondary school, and 183 (37.2%) completed university or college. Eight individuals left this question blank.

The mean number of years the couples had been together was 11.8 years (SD = 8.7), ranging from 1–39 years. Before the first session, study participants were also asked to identify their goals on a standard intake form. Three hundred sixty-six (73.5%) participants marked the goal of achieving a better relationship, whereas 118 (23.7%) sought clarification regarding whether the relationship should continue. Nine individuals (1.8%) indicated a goal of terminating the relationship in the best possible way, and another five (1.0%) marked "other" without elaboration. Two individuals did not answer. Three hundred six (61.2%) individuals were in a relationship in which both members marked the goal of achieving a better relationship, while 194 (38.8%) were in a relationship in which both had not marked the goal of achieving a better relationship.

Couples self-referred with a broad range of typical relationship problems, including communication difficulties, loss of feeling for partner, jealousy/infidelity, conflict, and coping with partner's physical or psychological problem. Diagnosis is not required nor a routine convention in this setting. The mean intake score on the Outcome Rating Scale (ORS; Miller, Duncan, Brown, Sparks, & Claud, 2003; discussed later) of the 500 participants was 19.01 (SD = 7.80), indicative of a clinical population and similar to distress levels of participants at other clinical sites (Miller & Duncan, 2004). Similarly, the mean marital satisfaction score on the Locke Wallace Marital Adjustment Test (LW; Locke & Wallace, 1959; discussed later) was 74.24 (SD = 24.55), indicative of a troubled relationship, well under the cutoff score of 100. The mean number of sessions completed was 4.32 (SD = 2.45). Sixty-nine couples attended two sessions (27.6%).

Follow-up participants. A total of 293 (58.6%) out of 500 individuals, representing 181 couples, responded to 6-month follow-up. In the follow-up sample, the couples were required to have data from both individuals for inclusion; that is, both members of the couple had attended at least two sessions of treatment and completed outcome measures (ORS and LW) for the first and follow-up evaluations. Since some couples completed just the ORS or LW at follow-up, the final sample size for the follow-up varied (115 couples for the ORS, and 101 couples for the LW). The mean ORS score at pretreatment was 19.83 (SD = 7.76). The mean marital satisfaction score on the LW at pretreatment was 79.91 (SD = 24.72). Although higher than the total sample, both measures indicate a pretreatment clinical population.

Subsample participants. To examine if clients' alliance scores would predict outcomes over early change and to discern couple alliance patterns, we selected clients who attended four sessions or more of couple therapy (n = 236 clients, 118 couples, and 19 therapists). The mean ORS score at pretreatment was 18.83 (SD = 7.53). The mean marital satisfaction score on the LW at pretreatment was 72.62 (SD = 24.11). Four sessions constituted the minimum number of sessions that would allow for both early change measurement and three data points for pattern analysis. The average number of sessions for this subsample was 6.17 (SD = 2.22).

Therapists

The couples were seen by 20 therapists (13 women and 7 men). Therapists worked at two family counseling agencies in Norway, 10 therapists from each agency. Ten were licensed psychologists, nine were licensed social workers, and one was a licensed psychiatric nurse. All therapists professed an eclectic orientation, using a variety of approaches—solution-focused, narrative, cognitive–behavioral, humanistic, and systemic—similar to those typically practiced in Norway family counseling agencies. The average age of the therapists was 44 years (SD = 12.6; range 26–61). The mean years of experience with couple therapy was 6.7 years (SD = 6.98; range, 0–19). The number of couples treated by each therapist ranged from four to 27, based on availability.

Measures

The Outcome Rating Scale (ORS). Psychological functioning and distress were assessed at the beginning of every session with the ORS (Miller & Duncan, 2004), but the analyses were derived from data collected at pre- and posttreatment and follow-up in the total sample, and at pretreatment, Session 3, and posttreatment in the subsample. The session that included the postassessment was variable, given that this was a naturalistic setting with no predetermined postsession. The ORS is a self-report instrument designed to measure client progress repeatedly (at the beginning of each session) throughout the course of therapy. The ORS is a four-item visual analog scale, reflecting key areas of client functioning: individually (personal well-being), interpersonally (family, couple, close relationships), socially (work, school, friendships), and overall (general sense of well-being). Clients put a mark on the line of each item nearest the pole that best describes their experience, and therapists score each 10-centimeter line using a centimeter ruler (each item is assigned a score ranging from 0 to 10). The scores are totaled, ranging from 0 to 40, with lower scores reflecting more distress.

Miller et al. (2003) reported that the internal consistency of the ORS was .93 and test–retest reliability was .66. In the current sample, the internal consistency of the ORS was .91. Concurrent validity of the ORS has been demonstrated in independent studies as adequate through correlates with the Outcome Questionnaire 45.2 (OQ; Lambert et al., 1996; r = .74, Campbell & Hemsley, 2009; r = .59, Miller et al., 2003). Using formulas developed by Jacobson & Truax (1991), Miller and Duncan (2004) analyzed the clinical and normative data for the ORS to provide cutoff scores for the reliable change index and clinically significant change. Based on a sample of 34,790 participants, clients who change in a positive or negative (deterioration) direction by at least 5 points are regarded as having made reliable change. This degree of change exceeds measurement error based on the reliability of the ORS and is one of the two criteria posited by Jacobson and Truax (1991) as indicative of clinically meaningful change. The second criterion requires movement from a score typical of a clinical population to one typical of a functional population. The cutoff on the ORS for marking the point at which a person's score is more likely to come

from a dysfunctional population than a nondysfunctional population is 25 (Miller et al., 2003).

Locke Wallace Marital Adjustment Test (LW; Locke & Wallace, 1959). The LW is a commonly used self-report measure of marital functioning. The LW is considered a reliable and valid measure of marital satisfaction and still relevant to clinical practice and research (Freeston & Plechaty, 1997). It is highly correlated with the often-used Dyadic Adjustment Scale ($r = .93$; Spanier, 1976). The LW cutoff score of 100, which differentiates satisfied from dissatisfied couples, is widely accepted (Christensen et al., 2004; Freeston & Plechaty, 1997). In the current study, the alpha for the LW was .75. The LW was administered at pretreatment and 6-month follow-up.

Session Rating Scale (SRS). The client's perspective of the alliance with the therapist was measured with the SRS (Duncan et al., 2003). The SRS was administered at the end of every session, but the analyses were derived from first and last sessions in the total sample and from first, third, and last sessions for the subsample. The SRS is also a four-item visual analog scale and is based on Bordin's (1979) classic delineation of the components of the alliance: the relational bond and the degree of agreement between the client and therapist about the goals and tasks of therapy. Clients place a mark on a 10-cm line nearest the pole that best describes their felt experience with their therapist. Specifically, the instructions of the SRS direct clients to rate their therapist on the following items: relationship with the therapist ("I felt heard, understood, and respected"), goals and topics ("We worked on or talked about what I wanted to work on or talk about"), the approach used in therapy ("The therapist's approach is a good fit for me"), and the overall rating of the session ("Overall, today's session was right for me"). The client's marks on the four items are measured with a centimeter ruler and totaled for a score ranging from 0 to 40. The cutoff score of 36 was derived from a sample of 15,000 clients, of whom only 24% scored below 36 and were at a statistically greater risk for a negative outcome (Miller & Duncan, 2004).

Hatcher and Barends's (1996) factor analysis of three popular alliance measures informed the construction of the SRS. They found that in addition to the general factor measured by all alliance scales (i.e., strength of the alliance), two other factors were predictive: *confident collaboration* and *the expression of negative feelings*. Confident collaboration speaks to the level of confidence that the client has that therapy and the therapist will be helpful. Although overlapping with Question 3 on the SRS (the fit of the therapist's approach), the fourth scale of the SRS directly addresses this factor and measures the client's view of the session ranging from "There was something missing in the session today" to "Overall, today's session was right for me." The other factor predictive beyond the general strength of the alliance is the client's freedom to voice negative feelings and reactions to the therapist. The SRS is intended to encourage clients to identify alliance problems so that the clinician may change to better fit client expectations.

Initial research has indicated the SRS generates reliable and valid scores. Duncan et al. (2003) found with a sample of 337 community mental health agency clients that the SRS had a coefficient alpha of .88 and was correlated with the Helping Alliance Questionnaire–II, $r = .48$ (HAQ-II; Luborsky et al., 1996). Test–retest reliabilities averaged .74 across the first six sessions with the SRS as opposed to .69 for the HAQ-II. Similar to other alliance measures, early SRS scores correlated significantly with outcome, $r = .27$. In an independent investigation, Campbell and Hemsley (2009) found the SRS was correlated with the Working Alliance Inventory, $r = .63$ (Horvath & Greenberg, 1989). In the current sample, the internal consistency of the SRS was .89. The SRS was developed to encourage clinicians to routinely assess and discuss the alliance with clients. The ORS and SRS [1] have a data base of over 200,000 administrations and have been used in two independent RCTs in which feedback was compared with TAU; one of these RCTs examined couple therapy (Anker et al., 2009; Reese, Norsworthy, & Rowlands, 2009).

Procedure

This was a naturalistic study conducted in community-based outpatient centers. Clients were invited to participate in a research study about improving the benefits of therapy. All participating clients gave their informed consent, and institutional review and approval was secured. Participant intake forms were assigned randomly and weekly to available therapist intake slots. Therapists could exchange one case for another if they felt uncomfortable with a couple's clinical presentation as depicted on the intake paperwork or had any previous nonclinical contact with a couple. Such an exchange happened 20 times over the course of the study, primarily because of previous nontherapy contact with the couple.

All therapists worked with couples utilizing the Partners for Change Outcome Management System (PCOMS; Duncan, 2010; Duncan, Miller, & Sparks, 2004: Miller, Duncan, Sorrell, & Brown, 2005) where therapists had access to alliance (SRS) and outcome (ORS) feedback from each person every session. PCOMS was based on Lambert and colleagues' (see Lambert, 2010) continuous assessment model using the OQ (Lambert et al., 1996), but there are differences beyond the measures. First, PCOMS is integrated into the ongoing psychotherapy process and routinely includes a transparent discussion of the feedback with the client (Duncan et al., 2004). Session-by-session interaction is focused by client feedback about the benefits or lack thereof of psychotherapy. Second, PCOMS assesses the therapeutic alliance every session and includes a discussion of the therapeutic relationship and any potential problems. Lambert's system includes alliance assessment only when there is a lack of progress.

All therapists attended 2 days of training (8 hr total) before the study and three 3-hr follow-up trainings during the investigation. Therapists were instructed to follow the general protocol outlined in the scoring and administration manual for the ORS and SRS (Miller & Duncan, 2004) as well as the transparent, collaborative process of monitoring outcome and the alliance with clients described in these authors' other publications (e.g., Duncan et al., 2004). Clients were administered the ORS by the therapists at the beginning of every session. Therapists were also instructed in the administration and use of the SRS (Duncan et al., 2003) to detect potential breaches in the alliance. Breaches were defined as a score less than 36 in total or less than 9 on any item of the SRS. Critical to the use of the SRS is that the client understands that it functions

[1] The Outcome Rating Scale and Session Rating Scale are free for individual clinician use and can be downloaded at www.heartandsoulofchange.com

to facilitate a conversation about the alliance between the client and the therapist. Also important for both client and therapist to understand is that the SRS carries no bad news but rather offers a way for therapists to improve and tailor services based on client preferences. Toward the end of every session, the SRS was administered to each client and scored, which allowed therapists to openly discuss any concerns and how the services may better fit client expectations. Therapists were instructed to candidly discuss any score less than 36 in total or less than 9 on any subscale. Although the procedures of this study strongly encouraged therapists to openly discuss the feedback with clients, the frequency or content of these interactions was not monitored.

Follow-up assessment was conducted 6 months after the last session. Each participant was mailed a packet containing a prepaid addressed envelope, the LW, the ORS, and other questions about their experiences in therapy. If no response was received within 3 weeks, another packet was sent.

Results

To address Hypotheses 1a, 1b, and 2, we developed three-level multilevel models (clients nested within couples who were nested within therapists) using the Actor–Partner Interdependence Analytical Model (APIM; Kashy & Kenny, 2000; Kivlighan, 2007). Use of the APIM avoids many of the complications of separate analyses for men and women; it models the mutual relationship between individuals, accounting for the interdependence between partner's scores. We conducted multilevel models utilizing the statistical package Hierarchical Linear Modeling (Version 6, or HLM6); Raudenbush, Bryk, Cheong, & Congdon, 2004). Table 1 provides the descriptive information for the changes in the ORS, SRS, and LW for men and women.

Full Sample: The Alliance and Outcome at Posttreatment and Follow-Up

We tested our first hypothesis—that individual and partner alliance scores after the first (Hypothesis 1a) and last (Hypothesis 1b) sessions would be positively related to therapy outcomes at posttreatment and follow-up—with the full sample.[2] Specifically, we predicted ORS at posttreatment by client gender, client and partner SRS scores (at first and last sessions), and four interaction effects between gender and SRS scores (at Level 1). We also controlled for client pretherapy functioning (ORS–pretreatment at Level 1) and number of sessions (at Level 2). As seen in Table 2, client and partner first-session alliance was not significantly related to ORS at posttreatment even when client and partner last-session alliance was not included in the model. Client ($d = 0.22$) and partner ($d = 0.30$) alliance scores at last session, however, were significant predictors of ORS at posttreatment, after the variance in the other variables was controlled (Table 2, first column). In other words, clients who reported a better alliance at the end of therapy had better therapy outcomes; client outcomes were also better when the partner had higher alliance scores with the therapist at the end of therapy. Men's alliance at last session was a stronger predictor of outcomes as compared with women's (i.e., significant Individual SRS–Last Session × Gender interaction, $d = 0.23$). The partial correlations between SRS scores at last session and ORS scores at posttreatment (after ORS at pretreatment was controlled) for men and women were .51 and .29, respectively.

Next, we examined the relationship between client and partner alliances with the therapist and follow-up scores. The predictor variables were the same as the previous model; however, we also controlled for ORS scores at last session in the prediction of ORS at follow-up. We did not have information about clients' LW at posttreatment, so in that model we only controlled for LW–pretreatment. The results, as seen in the second column in Table 2, revealed that only men's alliance was a significant predictor of ORS at follow-up, (i.e., significant Individual SRS–Last Session × Gender interaction, $d = 0.28$). The partial correlations between ORS at follow-up and SRS at last session (with ORS at pre- and posttreatment controlled) were .29 and .08 for men and women, respectively. Similarly, at follow-up, only men's alliance was a significant predictor of LW scores, (i.e., significant Individual SRS–Last Session × Gender interaction, $d = 0.39$). However, for men and women, we found that partner alliance also predicted LW scores at follow-up ($d = 0.35$), after controlling for the variance in the other variables. In other words, how individuals viewed the alliance with the therapist was predictive of partner LW score at follow-up. These results do not support Hypothesis 1a but do provide some support for Hypothesis 1b.

Subsample: Early Change and the Alliance

We hypothesized that individual and partner alliance scores would predict outcome over and above early treatment change (Hypothesis 2). We calculated early symptom change by subtracting client ORS–pretreatment scores from ORS–third-session scores. Since the ORS was completed prior to each session, this difference reflects the amount of change from two sessions of therapy. The average change was 5.82 ($d = .077$) for men and 5.67 ($d = 0.75$) for women. The mean early change from ORS–pretreatment score to ORS–second-session change was smaller, 3.57 ($d = 0.48$) for men and 1.83 ($d = 0.23$) for women. Given that the amount of change from pretherapy to third session for men and women was a medium-sized effect and above the reliable change index for the ORS (5-point change), this estimate better represents early symptom change.

In this model, the criterion variable was ORS–posttreatment, and the predictor variables were client and partner early symptom change and SRS at third session, gender, and four interaction effects between gender and client and partner early symptom change and SRS at third session (all at Level 1). We also controlled for ORS–pretherapy scores at Level 1. At Level 2, we included number of sessions attended by the couple. Table 3 shows the pretherapy/first, third, and posttreatment means and standard deviations for men and women on the ORS and SRS.

The results from the APIM multilevel model demonstrated that client third-session alliance was significantly related to ORS at posttreatment ($d = 0.25$), after we controlled for the other variables in the model, including early symptom change (see Table 4).

[2] Because Items 1 and 2 seem most similar to other measures of the alliance, whereas Items 3 and 4 are more interpretive of factor analytic findings of alliance scales, we reconducted the analyses using only Items 1 and 2. Findings paralleled those reported for the full scale.

Table 1
Means and Effect Sizes on the Outcome Rating Scale, Locke Wallace Marital Adjustment Test, and Session Rating Scale

	Men						Women					
	Pretreatment		Posttreatment		Follow-up		Pretreatment		Posttreatment		Follow-up	
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
SRS	32.83	5.02	34.82	4.91	—		34.12	4.25	35.82	4.80	—	
ORS	19.85	7.83	27.85	8.47	28.74[a]	7.43	18.12	7.68	26.11	8.96	29.22[a]	8.32
d ORS			1.02		1.14				1.04		1.45	
LW	80.02	26.08	—		91.35[b]	26.58	78.79	23.79	—		92.96[b]	26.52
d LW	—		—		0.43		—		—		0.60	

Note. $N = 250$ couples (500 individuals). SRS = Session Rating Scale; ORS = Outcome Rating Scale; LW = Locke Wallace Martial Adjustment Test.
[a] $N = 115$ couples (230 individuals). [b] $N = 101$ couples (202 individuals).

However, there was a significant gender interaction effect. Women's alliance score at third session was a stronger predictor of therapy outcomes than men's score. Partner alliance score at third session was not significantly related to therapy outcomes ($d = 0.13$), after the variance in the other variables was controlled. That is, we found that a client's outcome was not related to his or her partner's rating of the alliance with the therapist, after controlling for the variance of the other predictors. As expected, client early symptom change was a significant predictor of ORS scores at posttreatment ($d = 0.40$). However, partner early symptom change was not related to ORS scores at posttreatment ($d = 0.13$), suggesting that client outcome was relatively unrelated to partner experience of early symptom change. Additionally, the relationship between client and partner early symptom change scores and therapy outcome was consistent for women as compared with that for men (i.e., no significant interaction effects). Collectively, these results support our second hypothesis that the alliance predicted outcome over early change. Note that these results were consistent when last-session alliance was used as a predictor instead of third-session alliance.

Subsample: Alliance Patterns and Couple Outcomes

Lastly, to look at couple level outcomes specifically, we investigated whether different patterns of couple alliance development would differentiate couple outcomes (whether none, one, or both partners achieved reliable or clinically significant change; Hypothesis 3). To determine whether couples had discernable patterns in their alliance scores in the first, third, and last sessions, we conducted a couple-level cluster analysis. Using the couple-level file ($n = 118$), we entered women's and men's alliance score at first, third, and last sessions into a latent class (LC) cluster analysis with maximum likelihood estimate (Magidson & Vermunt, 2003). We conducted the LC cluster analysis using Latent Gold 4.5 software (Vermunt & Magidson, 2008).

Table 2
Fixed Effects From the Individual–Partner Interdependence Models: Predicting ORS–Posttreatment, ORS–Follow-Up, and LW–Follow-Up

	ORS–Posttreatment		ORS–Follow-up		LW–Follow-up	
Variable	γ	*SE*	γ	*SE*	γ	*SE*
Intercept	26.96***	0.47	27.66***		83.80***	1.45
Individual–Level 1						
Pretherapy functioning (ORS/LW)	0.38***	0.05	0.11	0.07	0.58***	0.06
Posttreatment–ORS	—		0.24**	0.07	—	
Gender	1.41***	0.32	0.01	1.26	−2.26	3.14
Individual SRS–First	−1.00	0.65	—		—	
Partner SRS–First	−0.90	0.55	—		—	
Individual SRS–Last	1.77**	0.64	0.61	0.48	−0.55	1.81
Partner SRS–Last	2.37***	0.48	0.66	0.46	8.17**	2.70
Individual SRS–First × Gender	0.74	0.96	—			
Partner SRS–First × Gender	0.70	0.84	—			
Individual SRS–Last × Gender	1.88*	0.95	2.11**	0.69	9.04**	3.09
Partner SRS–Last × Gender	−1.76	0.91	−1.55	0.93	−5.26	4.22
Couple–Level 2						
No. of sessions	0.26	0.14	0.18	0.15	−0.53	0.59

Note. Gender was coded 1 for men and 0 for women. *N*s *for* Outcome Rating Scale (ORS)–posttreatment = 500 individuals (250 couples) and 20 therapists. *N*s for ORS–Follow-Up = 230 individuals (115 couples). LW = Locke Wallace Marital Adjustment Test; SRS = Session Rating Scale; First = first session; Last = last session.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3
Subsample Pretreatment, Third Session, and Posttreatment Means and Effect Sizes on the Outcome Rating Scale and Session Rating Scale

	Men						Women					
	Pretreatment/first		Third		Posttreatment		Pretreatment/first		Third		Posttreatment	
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
SRS	33.20	4.97	34.40	5.13	35.97	3.80	34.04	4.13	35.45	4.78	36.52	3.35
ORS	19.92	7.49	25.74	9.32	29.87	8.10	17.73	7.57	23.40	9.76	27.71	8.71
d ORS			0.77		1.33				0.75		1.32	

Note. ORS = Outcome Rating Scale; SRS = Session Rating Scale.

We examined 2-, 3-, and 4-cluster models. On the basis of the interpretation of the models as well as model fit statistics—Bayesian information criterion (BIC)—we retained a 3-cluster model (i.e., lower BIC scores are preferable, and the BIC was 115 points lower for the three-cluster model compared with the two-cluster model; the BIC decreased only 11 points between the three-cluster and four-cluster models). Four couples were not identified by these clusters. Couple alliance scores by cluster are listed in Table 5. Cluster 1 or high linear couples started with high SRS scores at Session 1 and progressively scored higher until the end of therapy. Cluster 2 or moderate linear couples began with moderate SRS scores and similarly rated higher until the last session. Although high linear and moderate linear couples mirror one another, high linear couples (both men and women) had higher SRS scores throughout treatment. In contrast, Cluster 3 or low nonlinear couples had lower SRS scores over the course of therapy as compared with the other two. Women's SRS scores did not increase at Session 3 while men's SRS scores decreased; at the end of therapy, low nonlinear couples had SRS scores that were higher than those at Session 3, but men's SRS scores were consistent with those at Session 1 while women's scores were slightly higher.

Table 4
Fixed Effects for Individual–Partner Multilevel Modeling for Predicting ORS–Posttreatment by SRS at Session 3 and Early Symptom Change

	ORS–Posttreatment	
Variable	γ	*SE*
Intercept	28.88***	0.60
Individual–Level 1		
Pretreatment–ORS	0.59***	0.06
Gender	0.84	0.49
Individual SRS–Third	1.92**	0.84
Partner SRS–Third	−1.04	0.59
Individual Early Change	3.10***	0.87
Partner Early Change	1.02	0.71
Individual SRS–Third × Gender	−2.41*	1.13
Partner SRS–Third × Gender	2.33	1.23
Individual Early Change × Gender	1.27	1.33
Partner Early Change × Gender	−1.31	1.36
Couple–Level 2		
No. of sessions	0.44*	0.18

Note. The variance estimates for Levels 1–3 were 18.81, 23.55, and 1.06, respectively. Gender was coded 1 for men and 0 for women. $N = 118$ couples. ORS = Outcome Rating Scale; SRS = Session Rating Scale.
* $p < .05$. ** $p < .01$. *** $p < .001$.

The between-group differences for the three clusters in SRS scores were significant at $p < .001$ for all comparisons except moderate linear and high linear women's first-session alliance ($p = .03$), moderate linear and low nonlinear men's first-session alliance ($p = .02$), and moderate linear and high linear men's first-session alliance ($p = .08$). The within-group comparisons were significant except for low nonlinear women's first through third sessions ($p = .72$). All other within-group comparisons were significant ($p < .05$).

We conducted a chi-square analysis to examine whether the three alliance patterns differentially predicted couple therapy outcomes (i.e., both reached reliable or clinically significant change, only one partner reached reliable or clinically significant change, or neither partner reached reliable or clinically significant change). The results were statistically significant, $\chi^2(4, N = 114) = 14.48$, $p < .01$. As seen in Table 6, 77.1% of couples in the high linear cluster as compared with 45.5% of couples in the low nonlinear cluster achieved reliable or clinically significant change. In contrast, 31.8% of couples in the low nonlinear cluster as compared with 2.9% of couples in the high linear cluster attended a therapy in which neither partner reached reliable or clinically significant change.

Discussion

In the present study, we sought to (a) explore how client and partner alliances are related to therapy outcomes; (b) examine whether the alliance predicted outcome over and above early change; and (c) determine whether there were patterns in couple alliance scores that predicted couple outcomes. We found a significant relationship between the alliance and outcome for both individuals and their partners at the last session in the large naturalistic sample that included clients who only received two sessions. In a subsample containing couples who attended four or more sessions, we found that third- and last-session alliances predicted outcome over and above early change ($d = 0.25$). On this subsample, we also examined the relationship of couple alliance patterns and outcome and found a significant advantage for those couples who start with a high level of alliance scores and increase from there.

On a sample of 250 couples that included two session treatments, we found that first-session alliance ratings were not related

Table 5
Couples Alliance Patterns

	High linear (n = 35)				Moderate linear (n = 57)				Low nonlinear (n = 22)			
	Women		Men		Women		Men		Women		Men	
SRS	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
First session	35.83	3.08	35.87	4.14	33.51	3.92	34.05	3.57	28.52	5.69	31.41	4.18
Third session	37.99	1.53	38.68	1.53	34.80	3.46	36.40	2.24	27.92	5.91	27.89	5.24
Last session	39.44	0.50	39.42	0.57	35.80	1.87	37.09	1.73	31.34	4.73	31.39	3.24

Note. SRS = Session Rating Scale.

to outcome, but individual and partner last-session alliance ratings were associated with better outcomes at posttreatment.[3] That is, clients who reported a better alliance at the end of therapy had better outcomes, and client outcomes were also better when partner alliance scores were higher at the end of therapy. In the literature, first-session alliances have not typically been examined; later assessments, such as the third session, have been favored. We looked at first-session alliances in addition to third-session alliances (in the subsample) so that the nearly 28% of couples attending two sessions could be included, and therefore the sample would be more representative of routine practice. Given the temporal connection between the alliance at the last session and outcome, the finding of a significant relationship is perhaps not surprising because couples achieving good outcomes would likely report better alliances with their therapist.

The subsample of couples attending four or more session (n = 119) allowed a further examination of the relationship of the alliance to outcome after controlling for early change. Both third- and last-session alliances predicted outcome.[4] Two studies have indicated a similar significant association between third session alliance scores and outcome (Bourgeois et al., 1990; Johnson & Talitman, 1997). A divergent finding was reported by Knobloch-Fedders et al. (2007) who found an association between first-session alliances and improvement in marital distress, though Session-8 alliances did as well.

In the full sample, men's alliance scores at last session were a stronger predictor of therapy outcomes than women's scores. This pattern continued at follow-up for the ORS and LW. Symonds and Horvath (2004) similarly found that men's higher alliance scores were more linked to outcome than women's higher scores. Collectively, these findings add to the evidence of the significance of men's alliances in heterosexual couple therapy (e.g., Bourgeois et al., 1990; Brown & O'Leary, 2000; Knobloch-Fedders et al., 2007; Symonds & Horvath, 2004). Therapists may need to pay particular attention to ensuring men's connection to the process, not only in early stages but throughout treatment.

The gender-linked patterns emerging in the current study may be related to issues other than gender. In this study, women initiated the therapy 62% of the time, perhaps indicating a stronger commitment to the relationship or working to improve it. Therapists, therefore, may need to concentrate on developing strong alliances with either gender when the other partner initiated treatment. In other words, the differential role of men in our study and the literature in relation to alliance and outcome may be more a factor of commitment rather than gender. Further study is required to clarify the relationship between gender and client motivation in the formation of alliances.

Furthermore, in the subsample, women's third-session alliance scores were a stronger predictor than men's alliance scores. No explanation is readily apparent, but it may be that the alliance is interpreted differently when therapy is longer. Perhaps when couples invest longer term commitments to therapy, women's alliances emerge as the more critical, and when couples invest in a shorter term of therapy because of a reluctant partner (usually the man), men's alliance becomes more predictive. More research would provide better understanding of these differences.

Although the relationship between the alliance and outcome is largely considered fact, inconsistent findings in studies that attempt to parcel out early change have led researchers to question whether the alliance is a result of positive early changes in therapy rather than a process factor separate from symptom relief (Barber, 2009). We are not aware of any studies addressing this issue with couple therapy. Similar to Barber et al. (2000) and Klein et al. (2003), we found that client ratings of alliance at the third (and

[3] We also tested whether the regression coefficients for SRS at Session 1 and at posttreatment were statistically different in their prediction of outcome. Results revealed that individual SRS–posttreatment regression coefficient was not significantly different from individual SRS–pretreatment regression coefficient in the prediction of ORS–posttreatment, $\chi^2 = 0.86$, $p > .05$. However, partner SRS–posttreatment regression coefficient was significantly different from partner SRS–pretreatment regression coefficient in the prediction of ORS–posttreatment, $\chi^2 = 7.33$, $p < .01$. Additionally, the Individual SRS–Posttreatment × Gender interaction regression coefficient was also significantly different than Individual SRS–Pretreatment × Gender regression coefficient in the prediction of ORS–posttreatment, $\chi^2 = 4.32$, $p < .05$. Although the SRS–posttreatment was a significant predictor of ORS–posttreatment and the SRS–pretreatment was not, the SRS–posttreatment regression coefficient was not a stronger predictor of ORS–posttreatment than the SRS–pretreatment regression coefficient in all comparisons. These additional findings paint a more complex picture, and further research is needed to understand these effects.

[4] The difference between individual SRS at third-session regression coefficient and individual SRS–posttreatment regression coefficient in the prediction of ORS–posttreatment was statistically significant, $\chi^2 = 9.14$, $p < .01$, suggesting that SRS–posttreatment was a stronger predictor of therapy outcomes as compared with SRS at Session 3. Again, this may not be surprising given the temporal connection of administration of the measures. SRS–pretreatment did not predict ORS–posttreatment in the subsample as well.

Table 6
Couples Outcomes

	High linear (n = 35)		Moderate linear (n = 57)		Low nonlinear (n = 22)	
Member of couple who changed	%	*n*	%	*n*	%	*n*
Neither	2.9	1	12.3	7	31.8	7
One	20.0	7	36.8	21	22.7	5
Both	77.1	27	50.9	29	45.5	10

Note. Changed = achieved either reliable or clinically significant change. Percentages = percentage of couples who had *X* outcome within each cluster; n = number of couples within the cell. For instance, 77.1% of couples in the high linear cluster both changed, as compared with 45.5% of couples in the low nonlinear cluster.

last) session predicted outcome over and above early change. We defined early change as that exceeding the reliable change metric of the ORS, thereby surpassing changes attributable to error or chance and representing significant change in the distress level of the client. The relationship of the alliance to outcome did not appear to be merely a result of early symptom relief, even when that relief went beyond reliable change. While we found a relationship between the alliance and outcome beyond early change for clients, the influence of partner alliance with the therapist was not predictive of client outcomes. In other words, the systematic influence of partner alliance with the therapist on client outcomes appears to be minor after considering client alliance with the therapist and early symptom change. This finding differed from that for the full sample and may reflect the differences between the samples—that is, couples who attended four or more sessions and those who did not.

This study also addressed alliance patterns in couple therapy as they specifically related to couple outcomes. Couples who started with average or higher alliance scores and subsequently increased their alliance scores fared significantly better than the cluster of couples who did not. Couples improved the most in the cluster (high linear) that started with the highest alliance scores and continued to increase over the course of treatment. These results—combined with the findings that later sessions (third and last) were predictive while first-session alliances were not—point to the need for therapists to continually assess the alliance. These findings also suggest that the relative starting place of the alliance may not be as important as whether the alliance improves over the course of treatment. Optimism that lower early alliance scores can result in positive outcomes, given appropriate attention by the therapist, is also supported by Symonds and Horvath (2004). They found that increases in alliance ratings for both partners between Sessions 1 and 3 predicted outcomes for men. Determining whether there may be a critical alliance window between the first and third sessions or later requires further study.

Pinsof et al. (2008) asserted that training in couple therapy should encourage therapists to explore multiple aspects of the alliance, including individual perceptions of partner and couple alliances in conjoint treatment. The findings of our study support that assertion and suggest that training as well as therapy itself should focus on identification of different perspectives of the alliance so that therapists can directly address and rectify problems before they exert a negative influence on outcome. Therapists in training rarely have this opportunity and are most often supervised without any objective information about client responses to therapy (Sapyta, Riemer, & Bickman, 2005); exceptions, however, are emerging (Reese, Usher, et al., 2009; Sparks, Kisler, Adams, & Blumen, in press).

There are several limitations to this study. It is unclear if the alliance relationships found here can be shown with nonheterosexual couples or in more ethnically diverse service contexts or other countries. Although our study is one of the largest naturalistic investigations of the alliance in couple therapy, the use of only one outcome measure in the pre- through posttreatment analysis as well as only one alliance measure limits the conclusions that may be drawn. In addition, the instruments used were quite brief, potentially limiting our understanding of the impact of the alliance. We do not know if more extensive alliance and progress assessments would have given different results or whether other measures from clinician or observer perspectives would alter our findings. The depth and richness of information, both psychometrically and clinically, gleaned from longer alliance measures cannot be accomplished by a four-item alliance scale. This study was intentionally designed to more closely replicate what happens in routine clinical practice and was consequently limited in terms of the instruments chosen as well as the frequency of administration. Although the results on the LW at follow-up support some of our findings, the problems associated with the use of only one outcome measure in the primary analysis as well as the use of only one brief alliance measure remain.

A significant limitation is the potential influence of demand characteristics (Orne, 1962) or social desirability. Clients could have inflated their scores because the measures were completed in the therapist's presence, and clients knew therapists would likely discuss their meaning. This is more likely with the SRS (the alliance scale) than the ORS. Some clients do hide things from their therapist, but they are more likely to withhold a negative reaction to the therapist or session than to hide or misrepresent their level of distress (Farber, 2003). Conversely, having access to weekly feedback regarding the relationship may heighten the focus on the therapeutic alliance and promote active collaboration and disclosure of negative reactions. Although alliance scores tend to be positively skewed and clients tend to score alliance measures high regardless of whether they are in the presence of the therapist or not (e.g., Pesale & Hilsenroth, 2009), further research is needed that directly addresses demand characteristics. Our findings should be viewed with this important limitation in mind.

Similarly, responsiveness (Stiles, 1988, 1994) was of concern in this study because all therapists were trained to discuss alliance ratings and therefore were encouraged to intervene when alliance problems existed. It could be argued that since the alliance data was fed back to therapists during treatment, the usefulness of the alliance scores as a reliable predictor could have been reduced by inflation of the scores even beyond the typical skewed nature of alliance scores. Like Knobloch-Fedders et al. (2007) who faced a similar situation, we thought that client ratings of the alliance might be higher. But like Knobloch-Fedders et al., we found that this was not the case. Mean alliance ratings were lower than those in previous studies of the SRS (Duncan et al., 2003). Given the possibility of responsiveness—that feedback about the alliance influenced subsequent therapist intervention—the results seem to further support the importance of the alliance in couple therapy.

The findings of the current study further emphasize the well-traveled conclusions regarding the relationship of the therapeutic alliance to treatment outcome and hopefully add to the understanding of the complexities involved with alliance development in couple therapy. Our findings suggest that the alliance impacts outcome over and above symptom relief. Given the sparse and ambiguous nature of previous investigations of early change and other intervening variables, however, more research is needed before any conclusions can be drawn. This study also demonstrated the feasibility of continuous alliance assessment in routine clinical practice. With minimal training, therapists in the current study routinely monitored the alliance; it is possible that with more training (e.g., Crits-Christoph, Connolly Gibbons, Crits-Christoph, et al., 2006), continuous monitoring of the alliance could yield better outcomes.

A recent special section of *Psychotherapy: Theory, Research, Practice, Training* (Gelso, 2006) addressed the current status and future directions of the therapeutic alliance. Two recurrent themes emerged. First was the recognition that although the alliance has been robustly linked with outcome, the causal direction of this relationship has not been established (e.g., Crits-Christoph, Connolly Gibbons, & Hearson, 2006). And second was a call for routine alliance assessment. Castonguay, Constantino, and Holtforth, (2006) asserted: "The most obvious clinical implication of having viable measures of the alliance is that therapists should be using them—and they should especially ask their clients to fill them out" (p. 273). Crits-Christoph, Connolly Gibbons, Crits-Christoph, et al. (2006) also suggested that ongoing feedback should be given to clinicians on the quality of their alliances as both a research direction and as a way of directly improving client care. In addition, the Division 29 Task Force on Empirically Supported Relationships (Ackerman et al., 2001) has endorsed the routine assessment of the alliance. Given these recent endorsements, the evidence that supports the relationship between the alliance and outcome, and the possibility that the alliance influences outcome above early change, it appears time to routinely assess the alliance during treatment.

References

Ackerman, S. J., Benjamin, L. S., Beutler, L. E., Gelso, C. J., Goldfried, M. R., Hill, C., . . . Rainer, J. (2001). Empirically supported therapy relationships: Conclusions and recommendations of the Division 29 Task Force. *Psychotherapy: Theory, Research, Practice, Training, 38,* 495–497.

Anker, M. G., Duncan, B. L., & Sparks, J. A. (2009). Using client feedback to improve couple therapy outcomes: A randomized clinical trial in a naturalistic setting. *Journal of Consulting and Clinical Psychology, 77,* 693–704.

Barber, J. P. (2009). Towards a working through of some core conflicts in psychotherapy research. *Psychotherapy Research, 19,* 1–12.

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

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

Bourgeois, L., Sabourin, S., & Wright, J. (1990). Predictive validity of therapeutic alliance in group marital therapy. *Journal of Consulting and Clinical Psychology, 58,* 608–613.

Brown, P. D., & O'Leary, K. D. (2000). Therapeutic alliance: Predicting continuance and success in group treatment for spouse abuse. *Journal of Consulting and Clinical Psychology, 68,* 340–345.

Campbell, A., & Hemsley, S. (2009). Outcome Rating Scale and Session Rating Scale in psychological practice: Clinical utility of ultra-brief measures. *Clinical Psychologist, 13,* 1–9.

Castonguay, L. G., & Beutler, L. E. (2005). Common and unique principles of therapeutic change: What do we know and what do we need to know? In L. G. Castonguay & L. E. Beutler (Eds.), *Principles of therapeutic change that work* (pp. 353–369). New York, NY: Oxford University Press.

Castonguay, L. G., Constantino, M. J., & Holtforth, M. G. (2006). The working alliance: Where are we and where should we go? *Psychotherapy: Theory, Research, Practice, Training, 43,* 271–279.

Christensen, A., Atkins, D. C., Berns, S., Wheeler, J., Baucom, D., & Simpson, L. E. (2004). Traditional versus integrative behavioral couple therapy for significantly and chronically distressed married couples. *Journal of Consulting and Clinical Psychology, 72,* 176–191.

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 outcomes? Recommendations for future research on the alliance. *Psychotherapy: Theory, Research, Practice, Training, 43,* 280–285.

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

Duncan, B. L. (2010). *On becoming a better therapist.* Washington, DC: American Psychological Association.

Duncan, B. L., Miller, S. D., Reynolds, L., Sparks, J., Claud, D., Brown, J., & Johnson, L. D. (2003). The session rating scale: Psychometric properties of a "working" alliance scale. *Journal of Brief Therapy, 3*(1), 3–12.

Duncan, B. L., Miller, S. D., & Sparks, J. A. (2004). *The heroic client: A revolutionary way to improve effectiveness through client-directed, outcome-informed therapy* (Rev. ed.). San Francisco, CA: Jossey–Bass.

Farber, B. A. (2003). Patient self-disclosure: A review of the research. *Journal of Clinical Psychology, 59,* 589–600.

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.

Freeston, M. H., & Plechaty, M. (1997). Reconsideration of the Locke–Wallace Marital Adjustment Test: Is it still relevant for the 1990s. *Psychological Reports, 81,* 419–434.

Gelso, C. J. (2006). Working alliance: Current status and future directions: Editor's introduction. *Psychotherapy: Theory, Research, Practice, Training, 43,* 257.

Hatcher, R. L., & Barends, A. W. (1996). Patient's view of psychotherapy: Exploratory factor analysis of three alliance measures. *Journal of Consulting and Clinical Psychology, 64,* 1326–1336.

Holtzworth-Munroe, A., Jacobson, N. S., DeKlyen, M., & Whisman, M. A. (1989). Relationship between behavioral marital therapy outcome and process variables. *Journal of Consulting and Clinical Psychology, 57,* 658–662.

Horvath, A. O., & Bedi, R. P. (2002). The alliance. In J. C. Norcross (Ed.), *Psychotherapy relationships that work* (pp. 37–69). New York, NY: 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., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology, 38,* 139–149.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12–19.

Johnson, S. M., & Talitman, E. (1997). Predictors of success in emotionally focused marital therapy. *Journal of Marital & Family Therapy, 23,* 135–152.

Kashy, D. A., & Kenny, D. A. (2000). The analysis of data from dyads and groups. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 451–477). New York, NY: Cambridge University Press.

Kivlighan, D. M. (2007). Where is the relationship in research on the alliance? Two methods for analyzing dyadic data. *Journal of Counseling Psychology, 54,* 423–433.

Kivlighan, D. M., & Shaughnessy, P. (2000). Pattern of working alliance development. A typology of client's working ratings. *Journal of Counseling Psychology, 47,* 362–371.

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

Knobloch-Fedders, L. M., Pinsof, W. M., & Mann, B. J. (2004). The formation of the therapeutic alliance in couple therapy. *Family Process, 43,* 425–442.

Knobloch-Fedders, L. M., Pinsof, W. M., & Mann, B. J. (2007). Therapeutic alliance and treatment progress in couple psychotherapy. *Journal of Marital and Family Therapy, 33,* 245–257.

Lambert, M. (2010). "Yes, it is time for clinicians to routinely monitor treatment outcome." In B. L. Duncan, S. D. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart and soul of change. Delivering what works* (pp. 239–268). Washington, DC: American Psychological Association.

Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G. M., . . . Reisinger, C. (1996). *Administration and scoring manual for the OQ 45.2.* Stevenson, MD: American Professional Credentialing Services.

Locke, H. J., & Wallace, K. M. (1959). Short marital-adjustment and prediction tests: Their reliability and validity. *Marriage and Family Living, 21,* 251–255.

Luborsky, L., Barber, J., Siqueland, L., Johnson, S., Najavits, L., Frank, A., & Daley, D. (1996). The Revised Helping Alliance Questionnaire (HAQ-II): Psychometric properties. *The Journal of Psychotherapy Practice and Research, 5,* 260–271.

Magidson, J., & Vermunt, J. K. (2003). Comparing latent class factor analysis with the traditional approach in datamining. In H. Bozdogan (Ed.), *Statistical data mining and knowledge discovery* (pp. 373–383). Boca Raton, FL: Chapman & Hall/CRC.

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.

Miller, S. D., & Duncan, B. L. (2004). *The Outcome and Session Rating Scales: Administration and scoring manual.* Ft. Lauderdale, FL: Authors.

Miller, S. D., Duncan, B. L., Brown, J., Sparks, J., & Claud, D. (2003). The Outcome Rating Scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy, 2,* 91–100.

Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The partners for change outcome management system. *Journal of Clinical Psychology, 61,* 199–208.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17,* 776–783.

Patton, M. J., Kivlighan, D. M., & Multon, K. D. (1997). The Missouri Psychoanalytic Counseling Research Project: Relation of changes in counseling process to client outcomes. *Journal of Counseling Psychology, 44,* 189–208.

Pesale, F. P., & Hilsenroth, M. J. (2009). Patient and therapist perspectives on session depth in relation to technique during psychodynamic psychotherapy. *Psychotherapy: Theory, Research, Practice, Training, 46,* 390–396.

Pinsof, W. M., & Catherall, D. R. (1986). The integrative psychotherapy alliance. *Journal of Marital and Family Therapy, 12,* 132–151.

Pinsof, W. M., Zinbarg, R., & Knobloch-Fedders, L. M. (2008). Factorial and construct validity of the revised short form integrative psychotherapy alliance scales for family, couple, and individual therapy. *Family Process, 47,* 281–301.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. T., Jr. (2004). *HLM 6: Hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International.

Raytek, H. S., McCrady, B. S., Epstein, E. E., & Hirsch, L. S. (1999). Therapeutic alliance and the retention of couples in conjoint alcoholism treatment. *Addictive Behaviors, 24,* 317–330.

Reese, R., Norsworthy, L., & Rowlands, S. (2009). Does a continuous feedback model improve psychotherapy outcomes? *Psychotherapy, 46,* 418–431.

Reese, R. J., Usher, E. L., Bowman, D., Norsworthy, L., Halstead, J., Rowlands, S., & Chisholm, R. (2009). Using client feedback in psychotherapy training: An analysis of its influence on supervision and counselor self-efficacy. *Training and Education in Professional Psychology, 3,* 157–168.

Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: Theory, research, and practice. *Journal of Clinical Psychology/In Session, 61,* 145–153.

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family, 38,* 15–28.

Sparks, J. A., Kisler, T. S., Adams, J. F., & Blumen, D. G. (in press). Teaching accountability: Using feedback to train effective family therapists. *Journal of Marital and Family Therapy.*

Stiles, W. B. (1988). Psychotherapy process–outcome correlations may be misleading. *Psychotherapy: Theory, Research, Practice, Training, 25,* 27–35.

Stiles, W. B. (1994). Drugs, recipes, babies, bathwater, and psychotherapy process-outcome relations. *Journal of Consulting and Clinical Psychology, 62,* 955–959.

Stiles, W. B., Glick, M. J., Osatuke, K., Hardy, G. E., Sharpiro, D. A., Agnew-Davies, R., et al. (2004). Patterns of alliance development and the rupture-repair hypothesis: Are productive relationships U-shaped or V-shaped? *Journal of Counseling Psychology, 51,* 81–92.

Symonds, B. D., & Horvath, A. O. (2004). Optimizing the alliance in couple therapy. *Family Process, 43,* 443–455.

Vermunt, J. K., & Magidson, J. (2008). Latent GOLD 4.5. Belmont, MA: Statistical Innovations.

Received November 24, 2009
Revision received April 26, 2010
Accepted April 30, 2010 ■