

The Relationship Between Alliance and Outcome: Analysis of a Two-Person Perspective on Alliance and Session Outcome

Sigal Zilcha-Mano
University of Haifa

J. Christopher Muran and Clara Hungr
Adelphi University and Mount Sinai Beth Israel, New York,
New York

Catherine F. Eubanks
Yeshiva University and Mount Sinai Beth Israel, New York,
New York

Jeremy D. Safran
The New School for Social Research and Mount Sinai Beth
Israel, New York, New York

Arnold Winston
Mount Sinai Beth Israel, New York, New York

Objective: Better alliance is known to predict better psychotherapy outcomes, but the interdependent and interactive effects of both therapist- and patient-reported alliance levels have yet to be systematically investigated. **Method:** Using actor–partner interdependence model analysis the authors estimated actor, partner, and 2 types of interactive effects of alliance on session outcome in a sample of 241 patient–therapist dyads across 30 sessions of cognitive–behavioral and alliance-focused therapy. **Results:** Findings suggest that the most robust predictors of session outcome are within-treatment changes in patient reports of the alliance, which predict both patient and therapist report on outcome. Within-treatment changes in therapist reports of the alliance, as well as differences between patients and between therapists in their average ratings of alliance levels across treatment, predict outcome as reported by the specific individual. Although alliance was found to be a significant predictor of outcome in both treatments, for therapist-reported alliance and outcome it had a stronger effect in alliance-focused therapy than in cognitive–behavioral therapy. Additionally, dyads with the highest pooled level of alliance from both partners fared best on session outcome. **Conclusions:** The results are consistent with a 2-person perspective on psychotherapy, demonstrating the importance of considering the interdependent and interactive nature of both patient and therapist alliance levels on session outcome.

What is the public health significance of this article?

The study demonstrates that within-treatment changes in patient reports of the alliance may predict not only their own but also their therapists' reported subsequent session outcome. Within-treatment changes in therapist reports of the alliance as well as differences between patients and therapists in their average ratings of alliance levels across treatment, predict outcome as reported by the specific individual. Additionally, dyads with the highest pooled level of alliance from both partners fared best on session outcome.

Keywords: alliance, outcome, psychotherapy processes, 2-person psychology

The quality of the interpersonal interactions that occur between patients and therapists, and its importance for the success of treatment, have gained much empirical attention in recent decades.

One of the most promising directions of investigation concerns the *working alliance*, commonly defined as the emotional bond established in the therapeutic dyad and the agreement between patient

This article was published Online First April 7, 2016.

Sigal Zilcha-Mano, Department of Psychology, University of Haifa; J. Christopher Muran and Clara Hungr, Derner Institute of Advanced Psychological Studies, Adelphi University and Department of Psychiatry, Mount Sinai Beth Israel, New York, New York; Catherine F. Eubanks, Ferkauf Graduate School of Psychology, Yeshiva University and Department of Psychiatry, Mount Sinai Beth Israel; Jeremy D. Safran, Clinical Psychology Program, The New School for Social

Research and Department of Psychiatry, Mount Sinai Beth Israel; Arnold Winston, Department of Psychiatry, Mount Sinai Beth Israel.

The research was supported by Grant MH071768 from the National Institute of Mental Health (Principal Investigator: J. Christopher Muran).

Correspondence concerning this article should be addressed to Sigal Zilcha-Mano, Department of Psychology, University of Haifa, Mount Carmel, Haifa 31905, Israel. E-mail: sigalzil@gmail.com

and therapist concerning therapy goals and the tasks necessary to achieve them (Bordin, 1979; Hatcher & Barends, 2006). The strength of the working alliance is a consistent predictor of outcome in psychotherapy, with stronger alliance predicting better therapeutic outcomes (Horvath, Del Re, Flückiger, & Symonds, 2011). Recent studies have shown that this association is significant even after adjusting for temporal precedence between alliance and symptoms (Falkenström, Granström, & Holmqvist, 2013; Zilcha-Mano, 2016; Zilcha-Mano, Dinger, McCarthy, & Barber, 2014; Zilcha-Mano & Errázuriz, 2015).

Horvath et al. (2011) conducted a meta-analysis to evaluate the association between alliance and outcome, in which 112 of 175 independent effect sizes between alliance and outcome were measured according to the patient's point of view, and only 23 were based on the therapist's perspective (the rest were observer ratings). Across studies, therapists' mean alliance levels tended to be lower than those of their patients (Hatcher, Barends, Hansell, & Gutfreund, 1995; Horvath, 1994; Horvath & Symonds, 1991; Horvath & Luborsky, 1993), but the meta-analysis conducted by Horvath et al. (2011) suggests that both patient and therapist perspectives can predict psychotherapy outcome. Because previous studies rarely examined both the patient's and the therapist's alliance rating in the same analysis (Kivlighan, 2007), however, less is known about the unique contribution of each partner's report on the quality of the alliance to the alliance–outcome association. The few studies that have examined the unique contribution of each perspective to treatment outcome produced mixed findings, showing at times that (a) both patient and therapist alliance has a unique effect on outcome (Bachelor, 2013), (b) neither affects outcome (Knuuttila, Kuusisto, Saarnio, & Nummi, 2012), and (c) only one of them has a significant effect (Gullo, Lo Coco, & Gelso, 2012; Huddy, Reeder, Kontis, Wykes, & Stahl, 2012; Huppert et al., 2014; Marcus, Kashy, & Baldwin, 2009). But none of these studies examined the interdependence between patient and therapist alliance, which may explain some of this inconsistency.

Despite theoretical arguments in favor of a two-person approach to psychotherapy, in which both patient's and therapist's individual experiences and perceptions are seen to interact with one another to create therapeutic change (see Safran & Muran, 2000, for a review), most studies have not taken into account the interdependence between the two partners' views of the therapeutic dyad and its effect on treatment outcome. This approach ignores the fact that patients' and therapists' views of the alliance may influence each other (Rozarmin et al., 2008). This is also arguably a methodological shortcoming because the patient's and therapist's views on their relationship are not independent and are nested within the dyad (Krasikova & LeBreton, 2012). The actor–partner interdependence model (APIM) of dependencies within nested dyadic data was developed precisely to address this shortcoming (Kenny, 1996; Ledermann & Kenny, 2012) and is currently used in many fields of psychology (e.g., Cuperman & Ickes, 2009).

Despite its potential importance for evaluating the working alliance effect on outcomes (Kivlighan, 2007), only few studies have implemented the APIM method to study the alliance–outcome association in individual treatment (for treatment of couples and groups, see Anderson & Johnson, 2010; and Kivlighan, Lo Coco, & Gullo, 2015, respectively). The APIM differentiates

between two types of effects: *actor* and *partner*. The actor effect describes the relationship between a person's own rating of the predictor and the outcome (e.g., the patient's report on both alliance and outcome). The partner effect describes the relationship between a person's own rating of the predictor and the partner's rating of outcome (e.g., the patient's report on the alliance and the therapist's report on outcome). The APIM model enables us to take into account how both partners of the therapeutic dyad experience the therapeutic relationship and the interdependence between them and, thus, treats the alliance as a dyadic phenomenon.

Only two published studies have examined the alliance–outcome association in individual treatment using the APIM model, taking into account the effect of the interdependence between patient and therapist alliance on treatment outcome. Kivlighan (2007) reported that patients' and therapists' ratings of the alliance correlated significantly with their own ratings of several session–outcome measures, demonstrating an actor effect. But because the actor effect is susceptible to shared variance, perhaps more important is Kivlighan's (2007) finding that the therapists' ratings of the alliance also correlated significantly with their patients' ratings of session outcome, demonstrating a partner effect of differences between therapists. In a later publication on a different dataset, however, Kivlighan et al. (2014) reported the opposite partner effect. In this dataset, patient ratings of the alliance predicted therapist ratings of several session outcome measures, demonstrating a partner effect of differences between patients rather than between therapists as was found in Kivlighan (2007).

Similar inconsistencies were revealed in two studies that used APIM to investigate the real relationship (i.e., the personal, non-working aspect of the patient–therapist relationship; Gelso, 2014), a concept related to the alliance. Whereas Gelso et al. (2012) found that differences between patients in alliance rating predicted the therapists' ratings of outcome, Markin, Kivlighan, Gelso, Hummel, and Spiegel (2014) reported the opposite partner effect, in which differences between therapists in alliance rating predicted patients' ratings of outcome.

One possible explanation for these inconsistencies in studies that used the APIM model to examine the effect of the therapeutic relationship on outcome lies in the dynamic nature of alliance. It has been shown that alliance is not a fixed characteristic of a patient or a dyad but rather tends to develop and change over the course of treatment (Stiles & Goldsmith, 2010). The few pioneering studies that used APIM to account for the interdependence between patient–therapist alliance and outcome did not incorporate the developmental process of the alliance over the course of the treatment, untangling between- from within-individual effects.

To examine the effect of alliance as it unfolds over the course of treatment, one must disentangle the effects within and between individuals, as is true for the effect of any predictor on outcome over time (Curran & Bauer, 2011). The *within-individual* alliance effect on outcome reflects the association between time-specific changes in an individual's perception of alliance over time during treatment (e.g., improvement, no change, or deterioration relative to what is expected for the patient/dyad) and changes in outcome (e.g., improvement, no change, or deterioration of symptoms relative to what is expected for the patient/dyad) as reported by both the patient and the therapist (e.g., Falkenström et al., 2013; Hofart, Øktedalen, Langkaas, & Wampold, 2013; Zilcha-Mano &

Errázuriz, 2015). By contrast, the *between-individuals* alliance effect on outcome reflects the association between the individual's general tendency to report a better alliance and outcomes from their own or their partners' perspective (e.g., Baldwin & Imel, 2013; Baldwin, Wampold, & Imel, 2007; Curran & Bauer, 2011). To the best of our knowledge, no study to date has separated these two effects of within- and between-individuals alliance on outcome over the course of treatment, while also accounting for the interdependence between patients' and therapists' alliance effect on outcome. Therefore, it is still not known which are the most reliable predictors of outcome: differences between individuals, which we have termed *trait-like characteristics*, or time-specific changes over time within each specific individual, which we have termed *state-like characteristics* of alliance. The first aim of this study was thus to examine whether trait-like or state-like characteristics of the alliance are more robust in predicting outcome, beyond what may be explained by monosource bias (i.e., the same source, whether patient or therapist, rating both the predictor and the outcome).

When examining the interdependence between the effect of patient and therapist alliance on outcome, it is an open question whether alliance operates similarly across different treatments. Although the alliance is generally considered as a common factor, it may still operate differently in distinct treatment orientations (Ulvenes et al., 2012). Some theoretical orientations place greater emphasis on alliance as an active ingredient in treatment than others (Muran & Barber, 2010). Studies, however, generally failed to demonstrate the ability of treatment orientation to moderate the effect of alliance on outcome (Flückiger, Del Re, Wampold, Symonds, & Horvath, 2012, but see also Ulvenes et al., 2012). Although different treatments orientations were generally not found to produce different magnitudes of alliance effect on outcome, deliberate work on alliance with the aim of improving treatment outcome may enhance the effect of alliance on outcome. That is, although the alliance effect on outcome is common across treatments, its magnitude may be manipulated by deliberate work focusing on it. When examining the interdependence between patients and therapists it is of interest to examine whether a deliberate work on alliance with the aim of improving outcome affects the magnitude of the alliance–outcome association. In the present study, we used data from two treatments: one that focuses on constructive work with weakened alliance and negative therapeutic processes to improve treatment outcome (alliance-focused therapy [AFT]) and another that does not have such a focus (cognitive–behavioral therapy [CBT]).

In addition to the interdependence between patients and therapists, the importance of the two-person perspective lies also in the interactive effect between the two partners of the dyad. As reviewed above, several previous studies have demonstrated a potential effect of both patient- and therapist-reported alliance levels on their own and their partners' outcomes, but the *nature* of the interactive effects between partners' alliance levels has not yet been systematically investigated. This effect has been proven to be an important component of a dyad, both statistically (Malloy & Kenny, 1986) and conceptually. According to two-person approaches to psychotherapy, the therapist is considered a coparticipant in the treatment rather than someone who stands outside the interpersonal field and observes (Greenberg, 1995; Safran & Muran, 2000; Sullivan, 1954). Therefore, in addition to the impor-

tance of investigating the actor and partner effects of alliance on outcome, dyadic effects of the partners' alliance levels may also be important for understanding how alliance influences outcome. According to Kenny, Kashy, and Cook (2006), *dyadic interaction*, which is different from the common statistical concept of a multiplicative term, is important for understanding dyadic effects and can take several forms, each one valid as a test of “interaction effects,” including *difference* and *pooled dyadic* interactions.

In the context of therapist and patient alliance effect on outcome, both difference and pooled dyadic interaction effects have the potential to explain variance above and beyond the effects of either partner's alliance. Pooled interaction would indicate that partners' reports of outcome may be predicted by the sum of their own and their partners' rated alliance levels; in other words, dyads in which both partners report moderate alliance levels produce a similar outcome to dyads in which the patient reports low alliance and the therapist high alliance, and dyads in which the patient reports high and the therapist low alliance. At the same time, the difference interaction would indicate that it is the difference between the partners' alliance levels that can predict reported outcomes; in other words, dyads in which both partners report moderate alliance levels experience different treatment outcomes than do dyads in which the patient reports low alliance and the therapist high alliance, and in dyads in which the patient reports high and the therapist low alliance. The latter interactive effect has received some support in the literature (Bachelor, 2013; Kivlighan, 2007; Marmarosh & Kivlighan, 2012; Rozarmin et al., 2008), although the findings are inconsistent (Fitzpatrick, Iwakabe, & Stalikas, 2005; Meier & Donmall, 2006). Although only one of these potential effects received attention in previous studies (the difference interaction), both are theoretically tenable. Whereas dyads with the highest pooled level of alliance from both partners may fare best on session outcome, an inconsistency between the partners in the dyad may be indicative of relatively poor outcome.

Thus, the second aim of the present study was to examine two potential dyadic interactive effects of alliance on outcome: in the first interactive effect, the alliance is conceptualized as a pooled resource between partners, and we examine whether higher sums of alliance can predict a better outcome regardless of the contribution made by the perspective of each partner. In the second interactive effect, we examined whether differences in alliance levels between patient and therapist can serve as a sign of ineffective treatment. To the best of our knowledge, the present study is the first to examine these two potential effects together.

In sum, studies on the alliance–outcome association are only beginning to recognize the potential reciprocal effects that may exist within the therapeutic dyad. The few studies conducted in this area yielded promising but also inconsistent findings, presumably due to their reliance on only one snapshot of the alliance during treatment. Thus, they were not able to examine both the trait-like differences (between patients in therapist-reported and patient-reported alliance) and state-like changes (including strengthening, stability, and deterioration relative to the general tendency of the individual reported alliance) in the alliance over the course of treatment. The present study seeks to fill this gap in two ways: (a) it is the first study to examine the interdependence between patient- and therapist-reported alliance levels while disentangling differences between and within individuals alliance levels, for both therapist-rated and patient-rated alliance and outcome; (b) it is also

the first study to explore together two potential dyadic interactive effects (pooled vs. difference interactions) of both partners' alliance levels on session outcome over the course of treatment. With these two aims in mind, we investigated the ability of alliance to predict outcome using dyadic session-by-session monitoring of both alliance and session outcome across 30 sessions.

We used APIM (Kenny et al., 2006) for longitudinal data (Laurenceau & Bolger, 2012) to test patient and therapist alliance effects on their own session outcome (actor effects), their partners' session outcome (partner effects), and various dyadic interactive effects. We obtained the data from a randomized clinical trial (RCT) comparing the efficacy of CBT with that of AFT. Given that negotiating alliance is a main focus of AFT, we hypothesized that the effect of changes in alliance on outcome would be significantly stronger in this type of treatment than in CBT. Theoretically, alliance is perceived as a prominent mechanism of change in AFT, whereas in CBT other prominent mechanisms are assumed; therefore, the effect of changes in alliance on outcome is expected to be stronger in AFT, although we expect alliance to act as a mechanism of change in both treatments. Because the AFT condition focuses on training therapists to implement techniques designed specifically to enhance their ability to detect changes in the alliance and to work constructively with weakened alliance and with negative therapeutic processes to improve treatment outcome, we hypothesized that the stronger effect of alliance on outcome in AFT would be especially evident when focusing on the therapist's point of view (i.e., the therapist-actor effect). Following this hypothesis, we tested whether within-patient changes in patient-rated and therapist-rated alliance across treatment have stronger association with outcome in AFT than they do in CBT.

Method

Participants

Two hundred and 41 patients were randomly assigned as part of an RCT to one of two treatment conditions, CBT and AFT, at a large metropolitan medical center in New York City. The study was approved by the institutional review board of the relevant institution. Each patient had an equal chance of being assigned to each treatment condition. Patients were excluded from randomization for not meeting the following inclusion criteria: 18–65 years old and English fluency; or for meeting the following exclusion criteria (a) evidence of organic brain syndrome or mental retardation, (b) evidence of psychosis or need for hospitalization, (c) diagnosis of severe major depression or bipolar disorder, (d) evidence of active substance abuse, (f) evidence of active Axis III medical diagnosis, (g) history of violent behavior or impulse control problems, and (h) evidence of active suicidal behavior. Mean age was 42 ($SD = 13.54$) and 156 participants (64.7%) were female. One hundred and 79 (74.3%) were White, 6.7% Black, 5.8% Hispanic, and 13.2% chose the "other" category or did not answer this question. At intake, 49.8% met criteria for a primary diagnosis of mood disorder, 21.6% for anxiety disorders, and 4.6% for adjustment disorder; 46.1% met criteria for multiple Axis I diagnoses and 43.66% had a primary Axis II personality disorder. The most frequent personality disorders were avoidant (12%), obsessive–compulsive (10%), and not otherwise specified (20.7%). Of the patients, 58.9% were single, 19.5% married or

remarried, 14.5% divorced or separated, and 1.7% widowed. Four percent had some high-school education, 1.2% were high-school graduates, 14.9% had some college education, 38.6% college graduates, 7% had some postgraduate education, and 29.9% had graduate degrees.

Therapists

One hundred and 81 therapists participated in the study. They were clinical psychologists (7.1%), psychiatry residents (15.1%), and psychology interns and externs (77.4%). Mean clinical experience was 2.53 years ($SD = 2.87$), mean age was 31.24 ($SD = 4.17$), and 67.9% were women. Most of the therapists (76.1%) were White; and the rest were Latinos (7.7%), Asian (7.1%), Black (1.3%), or "other" (7.8%). The mean number of patients treated by each therapist in the current study was 1.32 ($SD = 0.59$, range: 1–4). Each therapist was randomized to conduct only one type of treatment in this RCT. Before being assigned a case, all trainees underwent an orientation seminar of six 1-hr lectures that provided an introduction to the theory, technique, and case formulation of the treatment modality to which they were randomized. Each trainee was then assigned a case screened for admission and began attending a weekly 90-min group supervision seminar. Each seminar was conducted by two senior supervisors with extensive experience in supervising the given treatment orientation. Therapists also received 1-hr weekly individual supervision by a senior supervisor during their first case. Those who were not licensed continued in individual supervision. Individual and group supervisions in both CBT and AFT made extensive use of videotaped sessions for feedback.

Treatments

Two treatment models were used in the study: AFT and CBT. Both models were manualized and designed to treat patients in a fixed, 30-session, one-session-per-week format. In the CBT condition (Turner & Muran, 1992), patients ($N = 108$) received treatment focusing on symptom reduction and schema change. The cognitive–behavioral strategies used included self-monitoring, cognitive restructuring, behavioral exercises, and experimentation. In the AFT condition (Muran & Safran, 2002; Safran & Muran, 2000; Safran, 2002) patients ($N = 133$) received treatment focusing on tracking alliance ruptures and engaging the patient in a process of metacommunication (e.g., facilitating communication about the communication process) to create awareness of the implicit negotiation between patient and therapist with regard to their respective needs. All the psychotherapy sessions were videotaped.

No significant differences were found at intake between the two conditions on either therapist or patient demographics, except for patient age, in which CBT patients were older ($M = 45.04$, $SD = 14.38$) than AFT patients ($M = 39.42$, $SD = 12.31$), $t(227) = 3.19$, $p = .002$. There was, however, a significant difference in patient pretreatment symptom severity between the two conditions, with the CBT patients showing more severe symptoms ($M = 0.99$, $SD = 0.65$) than AFT patients ($M = 0.79$, $SD = 0.54$), $t(235) = 2.49$, $p = .01$. Therefore, we conducted all analyses controlling for age and pretreatment symptom severity.

Measures

Working alliance. The quality of the working alliance was assessed with the Working Alliance Inventory (WAI; Tracey & Kokotovic, 1989) 12-item patient (WAI-P) and therapist (WAI-T) versions. Items were rated on a 7-point Likert scale, ranging from 1 (*never*) to 7 (*always*). In the present study, mean patient alliance rating ranged between 1.40 and 7, and therapist alliance rating ranged between 1.75 and 7. Studies support the psychometric properties of WAI measure (Elvins & Green, 2008). Research has shown that the WAI show a process of development over time for both patients (Falkenström et al., 2013; Zilcha-Mano & Errázuriz, 2015) and therapists (Zilcha-Mano et al., 2015), and that changes in alliance over time can significantly predict outcome for patient rating of alliance (Falkenström et al., 2013) and in some instances also for therapist rating (Zilcha-Mano et al., 2015). In the current study the internal reliability range across time points was .88–.94 for patients and .83–.93 for therapists.

Outcome. As a measure of session outcome, the one-item session outcome (Muran, Safran, Samstag, & Winston, 1992) measure was used repeatedly after each session for 30 weekly sessions. A single item was used to accommodate the time constraints of patients and therapists and to minimize self-report burnout. The one item was answered by both patients (“To what extent are your presenting problems resolved?”) and therapists (“To what extent are your patient’s presenting problems resolved?”) on a Likert scale, ranging from 1 (*not at all*) to 9 (*completely*). In the present study, patient session outcome rating ranged between 1 and 9 and therapist rating between 1 and 8. Each therapist and patient participating in the study received instructions from a research assistant trained to explain how to complete the session outcome item. The equivalence of one item over a full scale has been repeatedly demonstrated (e.g., Bergkvist & Ros-siter, 2007; Gardner, Cummings, Dunham, & Pierce, 1998; Robins, Hendin, & Trzesniewski, 2001). To further examine the validity of session outcome versus overall treatment outcome for the current data, we examined the association between the slope of change in overall treatment outcome from pretreatment to post-treatment, and the slope of change in session outcome, as reported by patients repeatedly over the course of treatment. Overall treatment outcome was assessed using the Global Severity Index (GSI) of the Symptom Checklist-90—Revised (SCL-90—R; Derogatis, 1983). In the current study the internal reliability of the GSI was .86 pretreatment and .84 posttreatment. Analysis yielded a significant moderate-to-high correlation, $r(108) = .58, p < .0001$.¹ The relatively high correlation between the session outcome measure and the overall treatment outcome measure used in the present study supports the validity of the session outcome as measured in the present study.

Procedure

After describing the study to the patients, written informed consent was obtained. Patients completed the SCL before starting treatment. Patient and therapist session outcome and working alliance ratings were collected session by session. Patients were informed that their therapists would not have access to their responses on these session measures. The mean length of treatment in the present study was 22.36 sessions ($SD = 10.17, Mdn = 29$). The post-SCL was administered at termination of treatment. Fur-

ther details on the design and procedures used are described elsewhere (Muran, 2002; Muran, Safran, Samstag, & Winston, 2005).

Treatment Fidelity

The observer-rated Beth Israel Fidelity Scale (BIFS) was used to assess the extent to which therapists conducted the treatments in accordance with the manuals. Studies have found that the BIFS showed sound psychometric properties, including adequate internal consistency, interrater reliability, and discriminant validity (Patton, Muran, Safran, Wachtel, & Winston, 1998; Santangelo, Safran, Muran, & Winston, 1994). In the present study, we used the following two subscales: (a) the AFT scale, 12 items developed to assess interventions associated with AFT; and (b) the CBT scale, 12 items developed to assess CBT interventions. Research assistants were trained to meet reliable standards (i.e., intraclass correlation $> .90$) in conducting this assessment. Eighty-two of the 241 patients participating in this study (34%) were randomly sampled to evaluate treatment fidelity (36 CBT and 46 AFT). One session was randomly selected from across the two treatments to assess early treatment fidelity on the BIFS (Sessions 3–7, $M = 4.77, SD = 0.91$). We conducted a series of *t* tests to examine differences on scale scores within each of the two treatments. Findings demonstrate that for each treatment condition therapists showed significantly higher ratings on the scales designed to measure the treatment model they were assigned to conduct (see Table 1).

Data Analysis

The data were hierarchically nested: sessions within individuals (patients or therapists), individuals within therapeutic dyads, and dyads within therapists. To account for this nonindependence of the data and to prevent inflation of the effects (Krull & MacKin-non, 2001; Laurenceau & Bolger, 2012), we used the SAS PROC MIXED procedure (SAS Institute, 2003), with Level 1 as the session level, Level 2 as the dyad level, and Level 3 as dyads of the same therapist. The session level was treated using the repeated option, whereas the dyad and therapist were assigned random effects. An advantage of the multilevel analyses lies in the flexibility in handling missing data (Gallop & Tasca, 2009).

To examine alliance and session outcome behavior over time, we evaluated the following trend models for each: linear, quadratic, linear in log of time, and stability over time either as fixed or random effects. We started with a model with only a fixed intercept and no random effects and added sequentially a random intercept, fixed effect of week, random effect of week, and a quadratic effect of week in therapy. Next, we examined the models with fixed and random linear effect of log of week. We used the log likelihood test and the Akaike information criterion to determine whether the inclusion of each term improved the model fit.

To disentangle between-patients and within-patient effects for both therapist and patient reports of alliance, we followed the

¹ The correlation is even higher, $r(108) = .63, p < .0001$, if we take into account the error resulting from deficits in the homogeneity of the Symptom Checklist (SCL), using a corrected *r* (dividing the correlation by the square root of the reliability of the SCL).

Table 1
Treatment Fidelity: Means, Standard Deviations, and *t* Tests of the Differences Between Fidelity Scores Within Each Treatment Condition

Condition and BIFS scale	<i>M</i> (<i>SD</i>)	<i>t</i>	<i>p</i>
CBT cases (<i>n</i> = 36)			
CBT scale	2.91 (0.53)	15.72	<.0001
AFT scale	1.47 (0.51)		
AFT cases (<i>n</i> = 46)			
CBT scale	1.40 (0.37)	-8.12	<.0001
AFT scale	2.22 (0.52)		

Note. BIFS = Beth Israel Fidelity Scale; CBT = cognitive behavioral therapy; AFT = alliance-focused therapy.

recommendations of Wang and Maxwell (2015) and centered the patient-reported and therapist-reported alliance within the individual patient's mean; we used the individual patient's mean for patient-reported and therapist-reported alliance for the between-patients effects. This procedure yielded independent coefficients for within-individual and between-individuals effects (Bolger & Laurenceau, 2013). Using this approach to disaggregate the within- and between-individuals components of alliance, we examined all alliance components simultaneously as predictors, in a combined model. We entered the pretreatment symptom severity (as measured by the SCL) into the analyses as a covariate.

The fact that our data were provided by dyads created an additional aspect of nonindependence to consider. To address this, we used the APIM (Kenny et al., 2006) for longitudinal data (Bolger & Laurenceau, 2013). APIM is a data-analytic approach designed specifically to test dyadic effects by simultaneously estimating actor effects (the effects of the actor's independent variable scores; e.g., their own alliance rating) on their own dependent variable score (e.g., their own session outcome rating) and partner effects (the effects of the partner's independent variable scores; e.g., the partner's alliance rating) on the actor's dependent variable score (e.g., the actor's session outcome rating). To lower the risk to the validity of the findings because of shared variance resulting from monosource bias (patient or therapist rating both the predictor and the outcome), we defined robust findings as instances in which the alliance as rated by one partner of the dyad (therapist or patient) can predict session outcome as rated by the same individual (an actor effect) as well as the other partner of the dyad (a partner effect). In other words, we focused on circumstances in which both actor and partner effects are significant.

The present data meet the requirements for APIM analysis (Bolger & Laurenceau, 2013; Kenny et al., 2006; for more information see the Appendix). Given our hypothesis that a stronger association is expected in AFT than in CBT between within-patient alliance, as reported by patients and therapists, and outcome, we tested the interactions between (a) within-patient changes in patient-reported and therapist-reported alliance and (b) treatment condition in predicting outcome.

The final model used was as follows:

$$Y_{ijk}(\text{outcome of session}_k \text{ for individual}_i[\text{patient or therapist}]$$

$$\text{in dyad}_j \text{ and therapist}_i) = \beta_{0ij} + \beta_1(\text{pre-treatment symptom severity}) + \beta_{r2}(\text{actor within-dyad level of alliance on}$$

$$\begin{aligned} & \text{session } k - 1) + \beta_{r3}(\text{actor between-dyad level of alliance}) \\ & + \beta_{r4}(\text{partner within-dyad level of alliance on session } k - 1) \\ & + \beta_{r5}(\text{partner between-dyad level of alliance}) \\ & + \beta_6(\text{role patient/therapist}) + e_{ijk} \end{aligned}$$

$$\beta_{0ij} = \beta_{0i} + e_{ij}$$

$$\beta_{0i} = \beta_0 + e_i$$

where e_{ijk} , e_{ij} , and e_i are independent normally distributed errors with zero means.

In the next step, to test our dyadic difference and pooled hypotheses, we followed the recommendations of Kenny et al. (2006). We calculated sum and absolute difference scores of actor and partner alliance levels both within and between patients, and we conducted the same mixed-model analyses with these variables instead of the centered crude scores. To test whether the sum or difference scores explain the outcome better than the parts do, we compared the fits of two models, the first one including the alliance sum or difference scores within and between patients and the second one including the patient-rated and therapist-rated alliance levels within and between patients. The two models were compared using the Bayesian information criterion (BIC) index.²

Results

Preliminary Analyses

We defined dropout as failure to complete the 30-week treatment protocol. On the basis of this definition, 39% of participants dropped out of treatment.³ Overall, therapists' average ratings of the alliance over time were significantly lower ($M = 4.98$, $SD = 0.69$) than patients' average ratings of the alliance ($M = 5.55$, $SD = 0.90$), $t(230) = 10.64$, $p < .0001$. The correlation between the two informants' average ratings was moderate in size and significant, $r(229) = .49$, $p < .0001$, suggesting that therapist and patient perspectives of alliance are related but not identical aspects of the alliance. The mean SCL reduction from pretreatment to posttreatment for treatment completers was -0.23 ($SD = 0.46$), which demonstrated significant change, $t(108) = 5.26$, $p < .0001$, without any significant differences between the AFT ($M = 0.20$, $SD = 0.44$) and CBT ($M = 0.27$, $SD = 0.49$) conditions, $t(107) = -0.79$, $p = .42$. Changes in patients' ratings of session outcome were not significantly different in the CBT ($M = 3.90$, $SD = 1.99$) and in the AFT condition ($M = 4.16$, $SD = 1.97$), $t(104) = -0.67$, $p = .49$. Significant differences were found in

² The Bayesian information criterion index was used to compare the fit of the two models rather than the significant level of specific coefficients. Therefore, the fact that the unique contribution of each coefficient is reduced when the two variables are placed together into the model (because of the correlation between them) is not expected to affect the fit of the model as a whole.

³ To examine a possible bias in our analyses as a result of patients who dropped out, we followed Gallop and Tasca's (2009) guidelines and classified available patient data into two patterns: patients who were followed up until week 30 (completers) and those who were not (dropouts). Next, we estimated the main model in our study adding the interactions of each term in the model with the pattern. An omnibus test comparing the two models revealed no significant increase in the model fit due to the addition of the interactions with pattern, $\chi^2(12) = 8.26$, $p = 0.76$.

changes in therapists' ratings of session outcome between the CBT ($M = 4.67$, $SD = 1.58$) and AFT conditions ($M = 3.76$, $SD = 1.65$), $t(127) = 3.15$, $p = .002$. We found no differences in changes in patients' alliance ratings from intake to posttreatment in the CBT ($M = 0.81$, $SD = 0.77$) and AFT ($M = 1.10$, $SD = 1.11$), $t(107) = -1.60$, $p = .11$, conditions and no significant differences in changes in therapists' alliance ratings in the CBT ($M = 0.94$, $SD = 0.87$) and AFT ($M = 1.20$, $SD = 0.98$), $t(133) = -1.58$, $p = .11$, conditions.

The simple correlations between patient-rated and therapist-rated alliance and patient-rated session outcome were .53 and .32, respectively, and the simple correlations between patient-rated and therapist-rated alliance and therapist-rated session outcome were .32 and .48, respectively. The aggregated session-level correlations using first-order differences of the correlation between patient-rated and therapist-rated alliance and patient-rated session outcome were .56 and .39, respectively. The correlations between patient-rated and therapist-rated alliance and therapist-rated session outcome were .36 and .40, respectively.

We compared the fits of several models of change over time for both alliance and session outcome, as reported by both patient and therapist. The model that was found to have the best fit is one with a fixed effect of linear trend in time and a random intercept for therapist- and patient-rated alliance for both CBT and AFT conditions, which is consistent with other reports on alliance development across treatment (Stiles & Goldsmith, 2010). For patient and therapist session outcome, the model that was found to have the best fit was the one with a fixed effect of log of time and a random intercept. This pattern of change across time is commonly found in psychotherapy research. We used this model in all analyses predicting session outcome.

Therapist's Random Effect

The estimated variance of the therapist's random effect in the three-level model was not significant ($S^2 = 0.15$, $p = .12$, $ICC = .06$). This finding indicates that the therapist's random effects did not contribute significantly to the variance of the session outcome. In other words, the difference in session outcome between therapists that is not explained by the covariates included in the model (patient- and therapist-rated alliance) was not significant. This unexplained difference in session outcome between therapists is represented by a random intercept for each therapist. The relatively small variance of the intercepts demonstrates their nonsignificant contribution to the model.

Dyad-Level Random Effect

The estimated variance of the dyad level random effect in the three-level model was significant ($S^2 = 0.65$, $p < .0001$, $ICC = .26$). This finding indicates that the random effect of the dyad contributed significantly to variance in session outcome. In other words, the difference in session outcome between dyads that is not explained by the covariates included in the model (patient- and therapist-rated alliance) was significant. This unexplained difference in session outcome between dyads is represented by a random intercept for each dyad. The relatively large variance of the intercepts demonstrates their significant contribution to the model.

Actor and Partner Effects

As can be seen in Table 2, analysis revealed significant actor effects (both within- and between-patients effects of both patient-rated and therapist-rated alliance) on session outcome. Both patients and therapists who reported time-specific increases from their mean alliance levels tended to report an increase in time-specific session outcome in the following session. Differences between patients in both patient- and therapist-rated alliance were also significantly associated with the individual's reported session outcome. In other words, patients and therapists who generally rated alliance as stronger across treatment also tended to report on better session outcome than did those who generally rated alliance as poorer. More important, analysis revealed significant partner effects for within-patient changes in patient-rated alliance levels on session outcome, where patients reporting time-specific increases from their mean alliance level had therapists reporting better session outcome.⁴

A significant interaction was found between changes in within-patient alliance, as rated by the therapist, and treatment condition on session outcome, as rated by the therapist: An increase in therapist-reported alliance predicted a greater increase in therapist-reported session outcome (actor effect) in the AFT than in the CBT condition ($B = -0.16$, $SE = 0.08$), $t(6576) = -1.89$, $p = .02$. In both treatments the effect of therapist-reported alliance on therapist-reported session outcome was significant (in AFT, $p < .0001$; and in CBT, $p = .01$). Controlling for treatment condition in all the other analyses did not affect the reported findings.

The Effect of Aggregated Levels of Patient- and Therapist-Rated Alliance on Treatment Outcome

Following the work of Crits-Christoph, Gibbons, Hamilton, Ring-Kurtz, and Gallop (2011), showing higher dependability when using aggregated measures of alliance, we repeated our analyses using aggregated measures of patient-rated and therapist-rated alliance (Sessions 2–10). Findings suggest that patient-rated alliance ($p = .04$) but not therapist-rated alliance ($p = .10$) was a significant predictor of treatment outcome, defined as change in SCL from pre- to posttreatment.

Difference and Additive Interactions Effects

The alliance sum score effect on outcome was significant both within patients ($B^5 = 1.43$, $SE = .03$), $t(8115) = 42.10$, $p < .0001$, and between patients ($B = 1.30$, $SE = .03$), $t(6792) = 41.45$, $p < .0001$, whereas the difference score effects on outcome were not,

⁴ Comparing the actor effect of within-patient changes in patient-rated alliance on patient-rated outcome and the partner effect of within-patient changes in patient-rated alliance on therapist-rated outcome revealed no significant differences between the actor and partner effects. We used standardized scores of the coefficients and compared them using a Wald test. We defined the estimated coefficient of the actor effect as b_1 and partner effect as b_2 , and we used the following statistic: $z = (b_1 - b_2) / \text{standard error}(b_1 - b_2)$, where the standard error is the square root of the following: $\text{variance}(b_1) + \text{variance}(b_2) + 2 \text{covariance}(b_1, b_2)$. The p value was calculated as follows: $2^*(1 - F(z))$, where F is the standard normal cumulative distribution function. The finding suggests no significant differences between the relevant actor and partner effects ($p = .54$).

⁵ All coefficients are unstandardized and therefore not comparable.

Table 2

Actor–Partner Interdependence Model for Between- and Within-Patient Alliance, as Rated by Therapists and Patients, and Session Outcome as Rated by Therapists and Patients

Effect	Coefficient	SE	df	t	p
Therapist session outcome intercept	3.78	.14	300	26.43	<.0001
Patient session outcome intercept	2.58	.14	300	18.03	<.0001
Actor affects					
Between-patients patient rating	.99	.09	255	10.43	<.0001
Between-patients therapist rating	.78	.12	248	6.48	<.0001
Within-patient patient rating	.16	.04	6602	3.52	<.0001
Within-patient therapist rating	.27	.04	6594	6.48	<.0001
Partner affects					
Between-patients patient rating	.10	.09	255	1.06	.28
Between-patients therapist rating	.07	.12	248	0.62	.53
Within-patient patient rating	.12	.04	6602	2.66	.007
Within-patient therapist rating	.03	.04	6594	0.72	.47

Note. An actor effect is the relationship between an individual's own alliance rating and their own outcome rating. A partner effect is the relationship between an individual's own alliance rating and the other individual's outcome rating. The *Bs* in the model are unstandardized and therefore not comparable.

$p \geq .15$. The interactions between the reporter role (patient vs. therapist) and the sum scores were not significant, $p \geq .43$. Therefore, we can assume that no differences exist in the sum of the alliances effect on between patient- and therapist-reported session outcome. To test whether the sum variable explains the outcome better than the parts do, we compared the fit of a model that includes the alliance sum variables with one that includes the patient- and therapist-rated alliance levels, using the BIC index. Findings suggest that the model with the sum variables showed a better fit than the one with patient-rated and therapist-rated alliance levels. Thus, our results support a pooled model in which alliance can be conceptualized as a pooled resource.⁶

Discussion

Although the therapeutic relationship between patient and therapist is often viewed as a dyadic construct, researchers examining the effect of alliance on outcome have rarely accounted for the dyadic aspects of this relationship (Kivlighan et al., 2014; Krasikova & LeBreton, 2012). To the best of our knowledge, no study to date has examined the effect of the development of alliance over the course of treatment on outcome as a dyadic concept. In this study, we focused on examining the effects of the interdependence between patient- and therapist-rated alliance levels across treatment and their interactive effect on outcome.

The first aim of this study was to examine the association between the interdependence of patient- and therapist-rated alliance levels and their session outcomes. We investigated the effect of changes in the alliance over the course of treatment on session outcome, adopting an analytic method that enabled us to treat the alliance as a dynamic construct. In a sample of 241 dyads, time-specific improvements from their average alliance level in the therapists' perception of the alliance over the course of 30 treatment sessions predicted their reports of the following session outcome. Patients and therapists who generally rated alliance as stronger across treatment also tended to report on better session outcomes than did those who generally rated the alliance as poorer. It is most important to note that time-specific improvements from their average alliance level in patients' perception of the alliance over the course of 30 treatment sessions emerged as a significant

predictor of both their own and their therapists' view of the following session outcome.

The findings concerning the first aim of the study demonstrate a between-patients alliance actor effect (as reported by both patients and therapists) that reflects the effect of a static, trait-like, overall characteristic of the alliance of a given patient on session outcome. The findings also demonstrate a within-patient alliance effect (for both patient- and therapist-rated alliance) that reflects the effect of time-specific changes in alliance over the course of treatment on session outcome. For both patients and therapists, state-like time-specific changes in alliance over the course of treatment affected their own perception of the following session outcome. However, only the state-like characteristic of alliance as rated by the patients had an effect not only on their own perception of the following session outcome but also on their therapists' perception of the following session outcome (namely, a significant partner effect).

The significant partner effect of within-patient changes in patient-rated alliance on therapist-rated outcome is a manifestation of the dyadic aspects of the alliance effect on outcome. The way in which patients perceived the alliance as it unfolds across the course of treatment predicts the therapists' perception of subsequent session outcome. In contrast to the actor effects, this finding cannot be interpreted as the result of shared variance because alliance and outcome were not rated by the same person. At the same time, within-patient changes in therapist-rated alliance were found to significantly predict therapist but not patient perception of subsequent session outcome. In other words, the way the therapist perceived the alliance as it unfolds across the course of treatment could not significantly predict the patient's perception of subsequent session outcome. This suggests that therapists should not expect their perception of changes in the alliance across treatment

⁶ Following Marmarosh and Kivlighan's (2012) recommendations, we repeated our analyses regarding the dyadic difference hypothesis, using a response surface analysis by polynomial regression instead of the absolute difference scores. The findings were consistent with the previous analyses and suggested no significant effect for agreement or for disagreement.

to affect patient outcome, unless the patients observe these changes themselves.

The present study is consistent with previous work recognizing the importance of the effect of both patient and therapist reports of alliance levels on treatment outcomes (Kivlighan, 2007; Kivlighan et al., 2014). It is also consistent with prior studies demonstrating the significant influence of both the actor and partner effects on the ability of the therapeutic relationship to predict outcome (Gelso et al., 2012; Kivlighan, 2007; Kivlighan et al., 2014; Markin et al., 2014). Our results contribute to the existing knowledge by demonstrating the importance of incorporating the development of alliance over the course of treatment in studies examining the interdependence between patient and therapist alliance and outcome. Of the handful of studies to date that have accounted for the interdependence between patients and therapists when examining the effect of the alliance on outcome, none have disentangled the trait-like (overall levels across treatment) and state-like (time-specific deviations from the individual's mean alliance level across treatment) characteristics of the individual's alliance. The findings of the present study can therefore explain inconsistencies found in previous studies, which did not separate differences between individuals from changes within individuals over the course of treatment, which may have been the source of the conflicting findings regarding the between-individuals alliance effect on outcomes (Kivlighan, 2007; Kivlighan et al., 2014). Our use of longitudinal data measuring the alliance–outcome association over the course of treatment is yet another important step in the process of defining a causal effect of alliance as a dyadic concept on outcome (Curran & Bauer, 2011).

We found further support in our analyses for the ability of patient-rated alliance to predict outcome based on aggregated measures of early alliance, which have the advantage of higher dependability (Crits-Christoph et al., 2011). These analyses revealed that patient-rated alliance was a significant predictor of treatment outcome (changes in SCL from pre- to posttreatment). Note that therapist-rated alliance was not a significant predictor of outcome. The inability of therapist-rated alliance to predict outcome is consistent with several previous studies (e.g., Zilcha-Mano et al., 2015). Given that in this study treatment outcome (changes from pre- to posttreatment based on the SCL) was reported only by patients, the finding is also consistent with the present finding based on an APIM model, that only actor effects not partner effects were found to be significant for therapists' self-report. In the present study, consistency across analyses is of great importance because of the complementary advantages of each method. Whereas analyses of aggregated early alliance have the advantage of greater dependability of the alliance measure and of reliable treatment outcome measure, analyses of the APIM model have the advantage of not relying exclusively on completers' data and of being able to disentangle within- and between-patients effects.

Findings also suggest differences between treatment conditions in alliance effect on outcome, with a greater effect of changes in therapist-rated alliance on therapist-rated outcome in the AFT than in the CBT condition. This differential effect can be understood based on the specific techniques used in AFT. The AFT condition includes the implementation of training procedures designed specifically to improve therapists' abilities to work constructively with weakened alliances and negative therapeutic process to improve therapists' abilities to repair

ruptures in the alliance when they occur. These techniques may have focused the therapist's attentiveness to changes in the alliance and facilitated the therapist's ability to put such changes to use in the interest of treatment success. This finding is consistent with previous literature demonstrating that some therapists consistently have better alliances and better outcomes than others (Baldwin & Imel, 2013; Baldwin, Wampold, & Imel, 2007; Wampold, 2001), as well as with studies demonstrating that therapists can be trained to enhance alliance and that such trainings may result in better outcomes, as manifest in the patients' quality of life (Crits-Christoph et al., 2006), interpersonal functioning (Safran et al., 2014), and lower dropout rate (Muran et al., 2005).

Regarding the second aim of the study, we explored the interactive effects of the therapeutic partners' alliance levels. In addition to testing the hypothesized actor and partner effects of alliance on outcome, we examined two conceptually distinct interactive effects between the partners' alliance levels: pooled and difference models. In a test of the potential interactive effects of the partners' alliance, our analysis supported a pooled model that considers alliance to be a dyadic resource from which both partners can benefit, to the extent that it is present. The findings regarding the nonsignificance of the difference model were replicated even when we used response surface analysis by polynomial regression instead of the absolute difference scores (see Footnote 1).

The significant pooled interactive effect may suggest that when either partner of the therapeutic dyad perceives the alliance as not being strong enough to enable their effective collaborative work together, it may be beneficial for the other partner to keep focusing on the strong elements of the alliance to keep alive the hope that the difficulties can be overcome and worked through, given the positive aspects of what the patient and therapist were able to form so far.

The findings relating to the second aim of the study are particularly interesting in light of previous studies examining the difference model in the alliance–outcome association. Several previous studies have found that the smaller the differences in alliance levels between patients and therapists, the better the outcome (Bachelor, 2013; Kivlighan, 2007; Rozarmin et al., 2008). Other studies, however, have found the opposite association, in which the greater the disagreement on alliance, the larger the improvement in symptoms is (Marmarosh & Kivlighan, 2012). Still others found no association between differences in alliance levels and outcome (Fitzpatrick et al., 2005) or dropout (Meier & Donmall, 2006).

Studies examining the effect of differences between patient- and therapist-reported alliance are based mostly on small sample sizes and differ extensively from one another in many aspects, such as the analytic method used, the time point in treatment in which alliance was examined, and the measure of treatment outcome. Significant moderators of the association between alliance differences and outcome may exist and explain the mixed findings. For example, almost each one of the cited studies used a different analytic method (e.g., polynomial regression and response surface analysis, difference scores, profile similarity correlations). Marmarosh and Kivlighan (2012) argued for major differences between these methods. But even when applying two different methods to the present data (polynomial regression and response surface analysis, and interactive effect of the differences), we found no significant effect for differences between patients and

therapists in alliance rating on outcome. Another potential moderator of the association between differences in alliance and outcome is the extent to which therapists are mindful of changes in alliance across treatment. In the present study, both treatments were conducted at a center that focuses on alliance research, which may have affected the therapists' treatment approach. Greater mindfulness of the alliance may have resulted in buffering any adverse effects of alliance disagreement. Some support for this post hoc explanation, arguing for the existence of moderators of the association between differences in alliance and outcome, may come from studies demonstrating significant moderators of the alliance–outcome association (e.g., Falkenström et al., 2013; Lorenzo-Luaces, DeRubeis, & Webb, 2014; Zilcha-Mano & Errázuriz, 2015). Future studies should test potential moderators of the association between differences in alliance and outcome and consider additional patterns of interactive effects, rather than focus strictly on the difference models that have been the subject of most of the research attention to date. Thus, the examination of alliance as a shared resource is another important contribution of the present study.

Overall, the study supports the importance of time-specific improvements in patient-reported alliance over time, relative to the general (average) alliance level reported by the patient, for the success of treatment, as evaluated by both patients and therapists. The findings are consistent with theoretical conceptualizations that stress the dynamic aspects of alliance (e.g., Safran & Muran, 2000) and with empirical studies describing the process of alliance development over the course of treatment (Stiles & Goldsmith, 2010). The findings support a view of alliance as a dyadic concept that emerges from a relationship and is embedded within it (Mitchell, 1995). Our results, therefore, suggest that the effect of alliance on outcome is multifaceted and best understood as dyadic and dynamic, rather than monadic and static.

Although we adopted a relatively cautious interpretation of the actor effects because of the risk of monosource bias (i.e., the fact that both predictor and product are reported by the same informant), several of the methods we used may further decrease this risk. First, we used lagged analyses, in which a temporal precedence is introduced between the predictor and the product. Second, disentangling between-patients effects from within-patient effects can also reduce any stable rater tendencies.

Our results should be considered in light of the limitations in our design, methods, and sample. The most important limitation of the present study is the use of a single item to evaluate session outcome. Although this practice is common in APIM applications, especially with a large number of repeated extensive observations recorded for both therapists and patients (in the present study, for many patients and therapists 30 observations were conducted, one per session), it has clear disadvantages. Although we found an association between session outcome and overall psychiatric symptom outcome (as measured by the SCL-90), which may suggest some inference between the two, the association is still far from perfect and concerns remain regarding the reliability and validity of a single-item measure of outcome. Thus, future studies should use a full scale to measure treatment outcome, as reported by both patients and therapists, after each session.

A second limitation of the study is the number of patients treated by each therapist. Although the number of therapists was relatively high compared to other studies, most therapists treated only two patients or fewer. The low number of patients treated by each therapist may have contributed to the nonsignificant random effect of therapists. Future studies should therefore use a higher therapist–patient ratio (Crits-Christoph et al., 2011). A third limitation has to do with the fact that following the recommendations of Kenny et al. (2006), we did not examine all the possible interactive effects between patient and therapist alliance and outcome. With adequate theoretical justification, future studies should examine additional interactions, such as synergistic influences (in which the effect of one individual's alliance on their own outcome is moderated by their partner's alliance levels), as well as cross-level interactions of between- and within-individual effects (Hofart et al., 2013). Fourth, although in the present study a linear development of between-sessions alliance was found across time, other patterns of alliance development may be observed (Kivlighan & Shaughnessy, 2000). Depending on the level of inspection and the methods used to measure patterns of alliance development (e.g., the WAI vs. the Rupture Resolution Rating System, Eubanks-Carter, Muran, & Safran, 2009), these patterns may include, for example, rupture-resolution patterns (Eubanks-Carter, Gorman, & Muran, 2012). Furthermore, different patterns may be revealed when looking at within-session changes in alliance as indications of ruptures, at between-session changes on a session-to-session basis (after the resolution of alliance ruptures may already occurred during the sessions), or at dysphoric discrete sessions.

A fifth limitation is that the present study did not focus on the direction of the disagreement between patients' and therapists' alliance levels. Future studies should pay attention to this and similar research questions (see Marmarosh & Kivlighan, 2012). Sixth, although patients and therapists did not have access to each other's ratings, many other factors may have contributed to their ratings, such as specific tendencies to answer self-report questions and demand characteristics. Although these factors may characterize each individual in the therapeutic dyad, they may also reflect interconnected elements of the dyad as an entity. Finally, the majority of the therapists in the present study were trainees (psychology interns and externs) who were randomized to conduct only one of two types of treatment; we may have obtained different results with a cohort of more experienced, senior therapists, or if the same therapists had conducted both types of treatment.

Overall, the current findings suggest that alliance has dyadic elements that should not be ignored. We found the most robust predictor of outcome to be time-specific deviations across treatment from the patient's general tendency to rate the alliance. Additionally, a pooled model in which both partners' alliance levels were combined to predict session outcome best evidence the interactive effects between patients' and therapists' alliance levels. We conclude that alliance may be best understood as a dyadic or relational construct: The results demonstrate the importance of considering the interdependent and interactive nature of the effect of patient- and therapist-rated alliance levels on session outcome.

References

- Anderson, S. R., & Johnson, L. N. (2010). A dyadic analysis of the between- and within-system alliances on distress. *Family Process, 49*, 220–235. <http://dx.doi.org/10.1111/j.1545-5300.2010.01319.x>
- Bachelor, A. (2013). Clients' and therapists' views of the therapeutic alliance: Similarities, differences and relationship to therapy outcome. *Clinical Psychology & Psychotherapy, 20*, 118–135. <http://dx.doi.org/10.1002/cpp.792>
- Baldwin, S. A., & Imel, Z. E. (2013). Therapist effects: Findings and methods. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed. pp. 258–297). New York, NY: John Wiley & Sons, Inc.
- Baldwin, S. A., Wampold, B. E., & Imel, Z. E. (2007). Untangling the alliance-outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology, 75*, 842–852. <http://dx.doi.org/10.1037/0022-006X.75.6.842>
- Bell, B. A., Schoeneberger, J. A., Morgan, G. B., Kromrey, J. D., & Ferron, J. M. (2010, April 11–14). Fundamental diagnostics for two-level mixed models: The SAS macro MIXED_DX. *Proceedings of the Annual SAS Global Forum*, Seattle, Washington.
- Bergkvist, L., & Rossiter, J. R. (2007). The predictive validity of multiple-item versus single-item measures of the same constructs. *Journal of Marketing Research, 44*, 175–184.
- Bolger, N., & Laurenceau, J. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York, NY: Guilford Press.
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research, Practice, Training, 16*, 252–260. <http://dx.doi.org/10.1037/h0085885>
- Crits-Christoph, P., Gibbons, M. B. C., 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., Gibbons, M. B. C., Hamilton, J., Ring-Kurtz, S., & Gallop, R. (2011). The dependability of alliance assessments: The alliance-outcome correlation is larger than you might think. *Journal of Consulting and Clinical Psychology, 79*, 267–278. <http://dx.doi.org/10.1037/a0023668>
- Cuperman, R., & Ickes, W. (2009). Big Five predictors of behavior and perceptions in initial dyadic interactions: Personality similarity helps extraverts and introverts, but hurts “disagreeables.” *Journal of Personality and Social Psychology, 97*, 667–684. <http://dx.doi.org/10.1037/a0015741>
- Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology, 62*, 583–619. <http://dx.doi.org/10.1146/annurev.psych.093008.100356>
- Derogatis, L. R. (1983). *SCL-90-R revised manual*. Baltimore, MD: John Hopkins School of Medicine.
- Elvins, R., & Green, J. (2008). The conceptualization and measurement of therapeutic alliance: An empirical review. *Clinical Psychology Review, 28*, 1167–1187. <http://dx.doi.org/10.1016/j.cpr.2008.04.002>
- Eubanks-Carter, C., Gorman, B. S., & Muran, J. C. (2012). Quantitative naturalistic methods for detecting change points in psychotherapy research: An illustration with alliance ruptures. *Psychotherapy Research, 22*, 621–637. <http://dx.doi.org/10.1080/10503307.2012.693772>
- Eubanks-Carter, C., Muran, J. C., & Safran, J. D. (2009). *Rupture resolution rating system (3RS): Manual*. New York, NY: Beth Israel Medical Center.
- Falkenström, F., Granström, F., & Holmqvist, R. (2013). Therapeutic alliance predicts symptomatic improvement session by session. *Journal of Counseling Psychology, 60*, 317–328. <http://dx.doi.org/10.1037/a0032258>
- Fitzpatrick, M. R., Iwakabe, S., & Stalikas, A. (2005). Perspective divergence in the working alliance. *Psychotherapy Research, 15*, 69–80. <http://dx.doi.org/10.1080/10503300512331327056>
- Flückiger, C., Del Re, A. C., Wampold, B. E., Symonds, D., & Horvath, A. O. (2012). How central is the alliance in psychotherapy? A multilevel longitudinal meta-analysis. *Journal of Counseling Psychology, 59*, 10–17.
- Gallop, R., & Tasca, G. A. (2009). Multilevel modeling of longitudinal data for psychotherapy researchers: II. The complexities. *Psychotherapy Research, 19*, 438–452. <http://dx.doi.org/10.1080/10503300902849475>
- Gardner, D. G., Cummings, L. L., Dunham, R. B., & Pierce, J. L. (1998). Single-item versus multiple-item measurement scales: An empirical comparison. *Educational and Psychological Measurement, 58*, 898–915. <http://dx.doi.org/10.1177/0013164498058006003>
- Gelso, C. (2014). A tripartite model of the therapeutic relationship: Theory, research, and practice. *Psychotherapy Research, 24*, 117–131.
- Gelso, C. J., Kivlighan, D. M., Jr., Busa-Knepp, J., Spiegel, E. B., Ain, S., Hummel, A. M., . . . Markin, R. D. (2012). The unfolding of the real relationship and the outcome of brief psychotherapy. *Journal of Counseling Psychology, 59*, 495–506. <http://dx.doi.org/10.1037/a0029838>
- Gonzalez, R., & Griffin, D. (1997). On the statistics of interdependence: Treating dyadic data with respect. *Handbook of Personal Relationships, 2*, 271–301.
- Greenberg, J. (1995). Psychoanalytic technique and the interactive matrix. *The Psychoanalytic Quarterly, 64*, 1–22.
- Gullo, S., Lo Coco, G., & Gelso, C. (2012). Early and later predictors of outcome in brief therapy: The role of real relationship. *Journal of Clinical Psychology, 68*, 614–619.
- Hatcher, R. L., & Barends, A. W. (2006). How a return to theory could help alliance research. *Psychotherapy: Theory, Research, Practice Training, 43*, 292–299. <http://dx.doi.org/10.1037/0033-3204.43.3.292>
- Hatcher, R. L., Barends, A., Hansell, J., & Gutfreund, M. J. (1995). Patients' and therapists' shared and unique views of the therapeutic alliance: An investigation using confirmatory factor analysis in a nested design. *Journal of Consulting and Clinical Psychology, 63*, 636–643. <http://dx.doi.org/10.1037/0022-006X.63.4.636>
- Hoffart, A., Økstedalen, T., Langkaas, T. F., & Wampold, B. E. (2013). Alliance and outcome in varying imagery procedures for PTSD: A study of within-person processes. *Journal of Counseling Psychology, 60*, 471–482. <http://dx.doi.org/10.1037/a0033604>
- Horvath, A. (1994). Research on the alliance. In A. Horvath & L. Greenberg (Eds.), *The working alliance: Theory, research and practice* (pp. 259–286). New York, NY: Wiley.
- Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy: Theory, Research, Practice, Training, 48*, 9–16. <http://dx.doi.org/10.1037/a0022186>
- Horvath, A. O., & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy. *Journal of Consulting and Clinical Psychology, 61*, 561–573. <http://dx.doi.org/10.1037/0022-006X.61.4.561>
- 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. <http://dx.doi.org/10.1037/0022-0167.38.2.139>
- Huddy, V., Reeder, C., Kontis, D., Wykes, T., & Stahl, D. (2012). The effect of working alliance on adherence and outcome in cognitive remediation therapy. *Journal of Nervous and Mental Disease, 200*, 614–619. <http://dx.doi.org/10.1097/NMD.0b013e31825bfc31>
- Huppert, J. D., Kivity, Y., Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2014). Therapist effects and the outcome-alliance correlation in cognitive behavioral therapy for panic disorder with agoraphobia. *Behaviour Research and Therapy, 52*, 26–34. <http://dx.doi.org/10.1016/j.brat.2013.11.001>

- Kashy, D. A., & Snyder, D. K. (1995). Measurement and data analytic issues in couples research. *Psychological Assessment*, 7, 338–348. <http://dx.doi.org/10.1037/1040-3590.7.3.338>
- Kenny, D. A. (1996). Models of non-independence in dyadic research. *Journal of Social and Personal Relationships*, 13, 279–294. <http://dx.doi.org/10.1177/0265407596132007>
- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York, NY: Guilford Press.
- Kivlighan, D. M., Jr. (2007). Where is the relationship in research on the alliance? Two methods for analyzing dyadic data. *Journal of Counseling Psychology*, 54, 423–433. <http://dx.doi.org/10.1037/0022-0167.54.4.423>
- Kivlighan, D. M., Jr., Gelso, C. J., Ain, S., Hummel, A. M., & Markin, R. D. (2015). The therapist, the client, and the real relationship: An actor-partner interdependence analysis of treatment outcome. *Journal of Counseling Psychology*, 62, 314–320. <http://dx.doi.org/10.1037/cou0000012>
- Kivlighan, D. M., Jr., Lo Coco, G., & Gullo, S. (2015). Is there a group effect? It depends on how you ask the question: Intraclass correlations for California Psychotherapy Alliance Scale-Group items. *Journal of Counseling Psychology*, 62, 73–78. <http://dx.doi.org/10.1037/cou0000046>
- Kivlighan, D. M., Jr., Marmarosh, C. L., & Hilsenroth, M. J. (2014). Client and therapist therapeutic alliance, session evaluation, and client reliable change: A moderated actor-partner interdependence model. *Journal of Counseling Psychology*, 61, 15–23.
- Kivlighan, D. M., Jr., & Shaughnessy, P. (2000). Patterns of working alliance development: A typology of client's working alliance ratings. *Journal of Counseling Psychology*, 47, 362.
- Knuuttila, V., Kuusisto, K., Saarnio, P., & Nummi, T. (2012). Early working alliance in outpatient substance abuse treatment: Predicting substance use frequency and client satisfaction. *Clinical Psychologist*, 16, 123–135. <http://dx.doi.org/10.1111/j.1742-9552.2012.00049.x>
- Krasikova, D. V., & LeBreton, J. M. (2012). Just the two of us: Misalignment of theory and methods in examining dyadic phenomena. *Journal of Applied Psychology*, 97, 739–757. <http://dx.doi.org/10.1037/a0027962>
- Krull, J. L., & MacKinnon, D. P. (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research*, 36, 249–277. http://dx.doi.org/10.1207/S15327906MBR3602_06
- Laurenceau, J., & Bolger, N. (2012). Analyzing diary and intensive longitudinal data from dyads. *Handbook of Research Methods for Studying Daily Life*, 1, 407–422.
- Ledermann, T., & Kenny, D. A. (2012). The common fate model for dyadic data: Variations of a theoretically important but underutilized model. *Journal of Family Psychology*, 26, 140–148. <http://dx.doi.org/10.1037/a0026624>
- Lorenzo-Luaces, L., DeRubeis, R. J., & Webb, C. A. (2014). Client characteristics as moderators of the relation between the therapeutic alliance and outcome in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, 82, 368–373. <http://dx.doi.org/10.1037/a0035994>
- Malloy, T. E., & Kenny, D. A. (1986). The social relations model: An integrative method for personality research. *Journal of Personality*, 54, 199–225. <http://dx.doi.org/10.1111/j.1467-6494.1986.tb00393.x>
- Marcus, D. K., Kashy, D. A., & Baldwin, S. A. (2009). Studying psychotherapy using the one-with-many design: The therapeutic alliance as an exemplar. *Journal of Counseling Psychology*, 56, 537–548. <http://dx.doi.org/10.1037/a0017291>
- Markin, R. D., Kivlighan, D. M., Jr., Gelso, C. J., Hummel, A. M., & Spiegel, E. B. (2014). Clients' and therapists' real relationship and session quality in brief therapy: An actor partner interdependence analysis. *Psychotherapy*, 51, 413–423. <http://dx.doi.org/10.1037/a0036069>
- Marmarosh, C. L., & Kivlighan, D. M., Jr. (2012). Relationships among client and counselor agreement about the working alliance, session evaluations, and change in client symptoms using response surface analysis. *Journal of Counseling Psychology*, 59, 352–367. <http://dx.doi.org/10.1037/a0028907>
- Meier, P. S., & Donmall, M. C. (2006). Differences in client and therapist views of the working alliance in drug treatment. *Journal of Substance Use*, 11, 73–80. <http://dx.doi.org/10.1080/14659890500137004>
- Mitchell, S. A. (1995). *Hope and dread in psychoanalysis*. New York, NY: Basic Books.
- Muran, J. C. (2002). A relational approach to understanding change: Plurality and contextualism in a psychotherapy research program. *Psychotherapy Research*, 12, 113–138. <http://dx.doi.org/10.1080/713664276>
- Muran, J. C., & Barber, J. P. (2010). *The therapeutic alliance: An evidence-based guide to practice*. New York, NY: Guilford Press.
- Muran, J. C., & Safran, J. D. (2002). A relational approach to psychotherapy: Resolving ruptures in the therapeutic alliance. In F. W. Kaslow (Ed.), *Comprehensive handbook of psychotherapy* (pp. 253–281). New York, NY: Wiley.
- Muran, J., Safran, J., Samstag, L., & Winston, A. (1992). *Patient and therapist postsession questionnaires, Version 1992*. New York, NY: Beth Israel Medical Center.
- Muran, J. C., Safran, J. D., Samstag, L. W., & Winston, A. (2005). Evaluating an alliance-focused treatment for personality disorders. *Psychotherapy: Theory, Research, Practice, Training*, 42, 532–545.
- Patton, M. J., Muran, J. C., Safran, J. D., Wachtel, P., & Winston, A. (1998). *Treatment fidelity measure for three brief psychotherapies*. New York: Beth Israel Center.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, 27, 151–161. <http://dx.doi.org/10.1177/0146167201272002>
- Rozarmin, E., Muran, J. C., Safran, J., Gorman, B., Nagy, J., & Winston, A. (2008). Subjective and intersubjective analyses of the therapeutic alliance in a brief relational therapy. *American Journal of Psychotherapy*, 62, 313–328.
- Safran, J. D. (2002). Brief relational psychoanalytic treatment. *Psychoanalytic Dialogues*, 12, 171–195. <http://dx.doi.org/10.1080/10481881209348661>
- Safran, J. D., & Muran, J. C. (2000). *Negotiating the therapeutic alliance: A relational treatment guide*. New York, NY: Guilford Press.
- Safran, J., Muran, J. C., Demaria, A., Boutwell, C., Eubanks-Carter, C., & Winston, A. (2014). Investigating the impact of alliance-focused training on interpersonal process and therapists' capacity for experiential reflection. *Psychotherapy Research*, 24, 269–285.
- Santangelo, C., Safran, J., Muran, J., & Winston, A. (1994). *Adherence to manualized brief psychotherapy treatments*. Poster presented at the Annual Meeting of the Society for Psychotherapy Research, Pittsburgh, PA.
- Stiles, W., & Goldsmith, J. (2010). The alliance over time. In J. C. Muran & J. P. Barber (Eds.), *The therapeutic alliance: An evidence-based guide to practice* (pp. 44–62). New York, NY: Guilford Press.
- Sullivan, H. S. (1954). *The psychiatric interview*. New York, NY: W. W. Norton & Company.
- Tracey, T. J., & Kokotovic, A. M. (1989). Factor structure of the working alliance inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 207–210. <http://dx.doi.org/10.1037/1040-3590.1.3.207>
- Turner, A. E., & Muran, J. C. (1992). *Cognitive-behavioral therapy for personality disorders: A treatment manual*. San Diego, CA: Social & Behavioral Documents.
- Ulvenes, P. G., Berggraf, L., Hoffart, A., Stiles, T. C., Svartberg, M., McCullough, L., & Wampold, B. E. (2012). Different processes for

- different therapies: Therapist actions, therapeutic bond, and outcome. *Psychotherapy*, 49, 291–302. <http://dx.doi.org/10.1037/a0027895>
- Wampold, B. E. (2001). Contextualizing psychotherapy as a healing practice: Culture, history, and methods. *Applied and Preventive Psychology*, 10, 69–86.
- Wang, L. P., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multi-level models. *Psychological Methods*, 20, 63–83. <http://dx.doi.org/10.1037/met0000030>
- Zilcha-Mano, S. (2016). New analytic strategies help answer the controversial question of whether alliance is therapeutic in itself. *World Psychiatry*, 15, 84–85. <http://dx.doi.org/10.1002/wps.20305>
- Zilcha-Mano, S., Dinger, U., McCarthy, K. S., & Barber, J. P. (2014). Does alliance predict symptoms throughout treatment, or is it the other way around? *Journal of Consulting and Clinical Psychology*, 82, 931–935. <http://dx.doi.org/10.1037/a0035141>
- Zilcha-Mano, S., & Errázuriz, P. (2015). One size does not fit all: Examining heterogeneity and identifying moderators of the alliance-outcome association. *Journal of Counseling Psychology*, 62, 579–591. <http://dx.doi.org/10.1037/cou0000103>
- Zilcha-Mano, S., Solomonov, N., Chui, H., McCarthy, K. S., Barrett, M. S., & Barber, J. P. (2015). Therapist-reported alliance: Is it really a predictor of outcome? *Journal of Counseling Psychology*, 62, 568–578. <http://dx.doi.org/10.1037/cou0000106>

Appendix

Central Requirements for Actor–Partner Interdependence Model (APIM) Analysis

The present data meets the requirements for actor–partner interdependence model (APIM) analysis (Bolger & Laurenceau, 2013; Kenny et al., 2006). APIM is based on the premise that one partner’s attributes and behaviors can affect the other partner’s outcome, which is a valid assumption in the context of the alliance–outcome association, as described in the introductory section. In our data, the partners of the dyad are distinguishable by role (therapist or patient), enabling us to use a meaningful variable to differentiate between them (Gonzalez & Griffin, 1997; Kashy & Snyder, 1995). When dyad members are distinguishable, the means, variances, and covariances for therapists and patients may be different. We modeled these possible differences by creating separate coefficients for patients and for therapists. We used hierarchical linear modeling to estimate actor and partner effects as they unfold over time. We treated patient and therapist data as nested within a dyad of $n = 2$. Patients and therapists completed versions of the same scales for alliance and session outcome. This is consistent with the APIM requirement that both partners use the same evaluation rating scales. As part of the process of testing whether the data meets the prerequisites for APIM analysis, we examined the assumption of normally distributed errors and homogeneity of variance for Level 1 and Level 2 (patients) and Level 3 (therapists; Raudenbush & Bryk, 2002) using SAS macro

mixed_ed (Bell, Schoeneberger, Morgan, Kromrey, & Ferron, 2010). We conducted a Shapiro–Wilk’s test to assess Level 1 normality of residuals, and we estimated multivariate skewness and kurtosis to test Level 2 and Level 3 multivariate normality. We estimated homogeneity of variance by Levene’s test for Level 1, Level 2, and Level 3 residuals.

Levene’s test of homogeneity of variance of Level 1 residuals confirmed the assumption about homogeneity of variance, that is, homoscedasticity was confirmed, $F(209, 6,570) = 1.15, p = .07$. Levene’s test of homogeneity of variance of Level 2 and Level 3 discretized absolute residuals also supported the validity of the model, $F(4, 205) = 1.51, p = .20$, and $F(4, 161) = 1.18, p = .32$, correspondingly. Tests for multivariate Level 2 skewness, $\chi^2(1) = 2.61, p = .11$, and kurtosis, $Z_{(\text{lower})} = -0.45, p = .67$; $Z_{(\text{upper})} = -0.83, p = .20$, indicated multivariate normality. Consistent with this, multivariate normality was also supported for Level 3 skewness, $\chi^2(1) = 1.83, p = .18$, and kurtosis, $Z_{(\text{lower})} = -0.67, p = .75$; $Z_{(\text{upper})} = -1.11, p = .13$.

Received July 6, 2015

Revision received February 4, 2016

Accepted February 4, 2016 ■

Instructions to Authors

For Instructions to Authors, please consult the January 2016 issue of the volume or visit www.apa.org/pubs/journals/ccp and click on the Instructions to Authors link in the Journal Info box.