

# Congruence of Therapeutic Bond Perceptions and Its Relation to Treatment Outcome: Within- and Between-Dyad Effects

Julian A. Rubel  
University of Trier

Eran Bar-Kalifa  
Bar Ilan University and Ben-Gurion University

Dana Atzil-Slonim  
Bar Ilan University

Sebastian Schmidt and Wolfgang Lutz  
University of Trier

**Objective:** The present study investigates the association between congruence of patients' and therapists' perceptions of the therapeutic bond and symptom improvement. **Method:** Bond congruence-outcome associations were examined on the within- and between-dyad level for 580 patients (mainly depression and anxiety) receiving cognitive-behavioral therapy. Symptom change was assessed on a session-to-session level as well as from pre- to posttreatment. For the between-dyad analyses, the truth and bias model was applied. For the within-dyad analyses, polynomial regression and response surface analysis were conducted. **Results:** On the between-dyad level, higher temporal congruence between patients' and therapists' bond ratings (i.e., their correlation) was associated with better treatment outcomes. Additionally, the average discrepancy between therapists' and patients' bond ratings showed a significant quadratic association with treatment outcome. A tendency for therapists to moderately rate the bond lower than their patients' showed lowest posttreatment symptom scores. On the within-dyad level, we found that when patients' and therapists' ratings were in "agreement," higher bond scores were associated with fewer next-session symptoms. For "disagreement," the results showed that if therapists rated the bond as weak, whereas their patients rated it as strong, higher subsequent symptom distress was observed than if patients rated the bond as weak and their therapists rated it as strong. **Conclusions:** The present study highlights the importance of therapists being vigilant to session-to-session changes in the therapeutic bond to adjust their interventions accordingly.

## What is the public health significance of this article?

This study found that dyads in which patients and therapists similarly perceived session-to-session changes in the therapeutic bond work more effectively to reduce patients' symptom impairment. Over the entire course of treatment, patients whose therapists reported moderately lower bond ratings than their patients achieved the best outcomes. On a session-level, when therapists and patients rated the bond in opposite directions, it seemed more detrimental when therapists perceived the bond as weak and patients as strong than vice versa.

**Keywords:** congruence, alliance, truth and bias model, response surface analysis, within and between-patient analysis

A preponderance of psychotherapy process-outcome research has dealt with the question whether the quality of the relationship between patient and therapist is associated with treatment outcome

(Crits-Christoph, Gibbons, & Mukherjee, 2013). Meta-analyses on this issue have found moderate but consistent associations for patients' as well as therapists' perceptions of the alliance (Flückiger, Del Re, Wampold, Symonds, & Horvath, 2012; Horvath, Del Re, Flückiger, & Symonds, 2011; Martin, Garske, & Davis, 2000).

While for many years alliance-outcome research was restricted to comparisons between patients (between-patient level), only recently, this line of research has been extended to comparisons within patients (within-patient level; e.g., Falkenström, Granström, & Holmqvist, 2013; Hoffart, Øktedalen, Langkaas, & Wampold, 2013; Rubel, Rosenbaum, & Lutz, 2017; Zilcha-Mano, Dinger, McCarthy, & Barber, 2014). On the within-patient level, comparisons are made between time points rather than patients. For instance, symptomatic impairment after sessions in which patients and/or therapists report better alliance quality is compared with impairment after sessions with worse alliance levels. Many have argued that these within-patient associations are better suited to

This article was published Online First February 1, 2018.

Julian A. Rubel, Department of Psychology, University of Trier; Eran Bar-Kalifa, Department of Psychology, Bar Ilan University, and Department of Psychology, Ben-Gurion University; Dana Atzil-Slonim, Department of Psychology, Bar Ilan University; Sebastian Schmidt and Wolfgang Lutz, Department of Psychology, University of Trier.

This work was supported by grants from the German Research Foundation (LU 660/10-1, LU 660/8-1). Julian A. Rubel and Eran Bar-Kalifa contributed equally to this article.

Correspondence concerning this article should be addressed to Julian A. Rubel, Clinical Psychology and Psychotherapy, Department of Psychology, University of Trier, D-54296 Trier, Germany. E-mail: rubel@uni-trier.de

answer questions about possible change mechanisms, as more possibly confounding variables can be excluded as alternative explanations for the alliance-outcome association (e.g., Curran & Bauer, 2011; Falkenström, Finkel, Sandell, Rubel, & Holmqvist, 2017; Hoffart, 2014; Wang & Maxwell, 2015). Consequently, recent research has focused on the within-patient alliance-outcome relationship. In accordance with the results from between-patient analyses, most of these investigations have found that a stronger alliance in one session was associated with lower symptom scores in the following session (e.g., Falkenström, Ekeblad, & Holmqvist, 2016; Falkenström et al., 2013; Hoffart et al., 2013; Kivlighan, Hill, Gelso, & Baumann, 2016; Rubel, Rosenbaum, & Lutz, 2017; Tasca & Lampard, 2012; Zilcha-Mano et al., 2016; Zilcha-Mano, et al., 2015). Only few studies did not replicate this finding (e.g., Sasso, Strunk, Braun, DeRubeis, & Brotman, 2016).

Although both the patients' and therapists' perception of the alliance have been shown to be associated with treatment outcome on a within- and between-patient level, these perceptions are often only moderately related to one another (e.g., Tryon, Blackwell, & Hammel, 2007). Studies investigating alliance congruence have typically operationalized this relation as either the correlation between patients' and therapists' alliance ratings over two or more sessions or as the average difference between these ratings. In a meta-analysis reviewing 32 studies applying a correlational definition of congruence, Tryon et al. (2007) found a moderate aggregated effect of  $r = .36$ . Moreover, Tryon et al. (2007) found 44 studies in which the difference score between patients' and therapists' alliance ratings was used as a measure of congruence. On average, therapists rated the alliance lower than their patients with a standardized mean difference of  $d = 0.63$ . This moderate correlational association and substantial difference in average alliance levels lends support to the view that because of their different roles in the therapeutic process (Bachelor, 2013; Hartmann, Joos, Orlinsky, & Zeeck, 2015) and their differential prior experiences with the therapeutic setting (Fenton, Cecero, Nich, Frankforter, & Carroll, 2001; Tryon et al., 2007) patients and therapists take somewhat different perspectives toward their relationship.

However, these average associations have been shown to vary between and within dyads (e.g., Atzil-Slonim et al., 2015; Bachelor & Salame, 2000; Marmarosh & Kivlighan, 2012). On the between-dyad level, some dyads show high correlations between patient and therapist perceptions of the alliance, whereas others show low correlations. Similarly, some dyads show small differences in their average patient and therapist alliance ratings, whereas others show large differences; with either the patient or therapist rating the alliance lower than the other.

On a within-dyad level, patient and therapist ratings are in agreement in some sessions, whereas in other sessions they are in disagreement; again with either the patient or therapist rating the alliance lower than the other.

From a clinical perspective, it has been argued that a shared view on the therapeutic relationship may be an important characteristic of a successful patient-therapist dyad (e.g., Hartmann et al., 2015; Marmarosh & Kivlighan, 2012; Pepinsky & Karst, 1964). In cognitive-behavioral therapy (CBT) the role of the therapeutic alliance has been compared to the role of anesthesia during surgery (Castonguay, Constantion, McAleavey, & Goldfried, 2010; Raue & Goldfried, 1994). As such, the alliance can be seen as a necessary change factor in CBT, without which a successful application

of specific techniques is unlikely. Following that, the quality of the alliance is an important indicator for therapists, guiding their clinical decision-making regarding which interventions should be applied and how to do so. Therefore, matching therapist and patient alliance perceptions should facilitate successful therapeutic work. If therapists perceive the alliance differently than their patients, their interventions may plausibly be inappropriately attuned to the expectations and needs of their patients. For example, if the therapist perceives the alliance quality as low and the patient perceives it as high, the therapist might falsely refrain from interventions, which require a certain amount of trust and respect. Instead, therapists may apply interventions designed to strengthen the alliance although the patient may be already expecting more symptom-oriented interventions. If, in contrast, the therapist experiences a good alliance quality and the patient does not, the patient may feel overwhelmed by the interventions applied by the therapist because of the lack of mutual trust. In both cases, the conflicting alliance perceptions may result in a less effective therapy process.

Given its potential importance, as described above, the agreement between therapist and patient alliance ratings and its association with treatment outcome has generated considerable research interest of late. However, previous research on alliance congruence is limited in several ways: First, the use of difference scores has been criticized for various reasons (e.g., Marmarosh & Kivlighan, 2012). Second, studies typically utilized only one or just a few time points to assess agreement in the therapeutic relationship. Third, there is scant research on the association between congruence and treatment outcome. Fourth, no study so far has taken into account the different levels at which congruence-outcome associations take place; namely, the within- and between-dyad level.

To address these methodological concerns, more sophisticated multilevel regression-based methods have been introduced to retrieve different congruence indicators and to investigate their associations with treatment outcome: the truth and bias model (T&B; West & Kenny, 2011) and polynomial regression response surface analysis (RSA; Edwards & Parry, 1993). Both models have recently gained importance in the study of alliance agreement and its effects in counseling and psychotherapy (e.g., Atzil-Slonim et al., 2015; Compare, Tasca, Lo Coco, & Kivlighan, 2016; Marmarosh & Kivlighan, 2012; Zilcha-Mano, Snyder, & Silberschatz, 2016).

The T&B model (West & Kenny, 2011) allows the simultaneous estimation of two indicators of alliance overlap: The *temporal congruence* of the perspectives and the average *directional discrepancy*. With regard to alliance research, congruence reflects the correlation between patients' and therapists' ratings over the course of treatment. That is, high congruence refers to mutual fluctuations in alliance ratings, while low congruence represents asymmetrical changes. Discrepancy is a measure of the average amount the two perspectives differ over the course of treatment (cf. Atzil-Slonim et al., 2015).

The alternative method, namely polynomial regression and response surface analysis (RSA), is especially apt to investigate the association between two predictor variables, such as patient and therapist alliance ratings, and a dependent variable (e.g., next session symptom impairment). Additionally, it can be used to examine whether agreement has a beneficial effect on the criterion variable. While results from RSA provide more information on the effects of absolute alliance ratings in the case of agreement (patient

and therapist ratings match) and in the case of disagreement (patient and therapist ratings are opposite to each other), the T&B model analysis complements these results with global indicators of the average difference between the two perspectives and their congruence. Therefore, in the present study, both models were applied to obtain a comprehensive understanding of the association between agreement of alliance experiences and outcome.

So far, results on the association between agreement in alliance ratings and treatment outcome are mixed. For the analysis of this association, the differentiation of within- and between-patient/dyad comparisons described above is again pivotal. Most of the previous studies have focused on between-patient/dyad comparisons. Some of these studies have reported finding evidence of a significant association between alliance agreement and outcome (Fjermstad et al., 2016; Rozmarin et al., 2008). For example, Rozmarin et al. (2008) investigated the correlation between patients' and therapists' weekly assessments of the 12 item version of the working alliance inventory (WAI-S; Tracey & Kokotovic, 1989) as an indicator of congruence. They found this correlation to be significantly associated with treatment outcome in a diagnostically heterogeneous sample of 22 patients treated with 30 sessions of brief relational therapy. Dyads with higher correlations over the course of treatment showed better outcomes at the end of therapy. This association was even twice as strong as the respective associations of patients' and therapists' alliance ratings with outcome.

On the other hand, other studies did not find such an association (Compare et al., 2016; Fitzpatrick, Iwakabe, & Stalikas, 2005; Langhoff, Baer, Zubraegel, & Linden, 2008; Marmarosh & Kivlighan, 2012; Sinclair & Holmqvist, 2013; Zandberg, Skriner, & Chu, 2015). For example, Sinclair and Holmqvist (2013) conceptualized agreement as the average absolute difference between patients' and therapists' weekly WAI-S ratings and found no association with psychological distress after treatment in a sample of 700 patients with mixed diagnoses treated by one of 14 different therapy orientations in primary care centers in two Swedish regions.

On the within-dyad level only, Zilcha-Mano, Snyder, & Silberschatz, (2016) investigated the associations between alliance agreement measured with the WAI-S and symptoms measured with the OQ. For 127 diagnostically heterogeneous patient-therapist dyads, alliance agreement in one session was used as a predictor of symptoms 1 month later during the course of cognitive-relational psychodynamic treatment. Applying response surface analysis, Zilcha-Mano, Snyder, and Silberschatz (2016) found that patient and therapist ratings had a combined association with subsequent symptoms. One month later, symptom levels were lower when patients' and therapists' perceptions of the alliance were in agreement and (the alliance was) good than when they were in agreement and (the alliance was) moderate or poor. Interestingly, when both the therapist and the patient agreed that the alliance was poor, the patient's symptom level 1 month later was lower than when the two were in agreement and (the alliance was) moderate. The authors conclude that this finding may hint at the potential of alliance ruptures, which are recognized by the patient and the therapist (Zilcha-Mano, Snyder, & Silberschatz, 2016). With regard to disagreement, the authors neither found that the amount of disagreement had any association with symptom impairment, nor did one of the two perspectives contributed more or less to subsequent symptom distress. This last result is in disagree-

ment with the common finding of several alliance-outcome studies that patients' alliance ratings are stronger associated with patient-rated symptoms than therapists' alliance ratings (e.g., Zilcha-Mano et al., 2016).

### The Present Study

Given the vast heterogeneity between samples in previous research (e.g., applied methods, number of alliance measurements, treatment orientations, treatment length, treatment targets, modalities, and alliance questionnaires) it is quite difficult to resolve from where inconsistencies in the literature may arise. The present study is the first to apply two state-of-the-art methods of congruence research, the truth and bias model and response surface analysis, to a longitudinal dataset of session-to-session alliance measures to be able to determine between- and within-dyad congruence-outcome associations in individual CBT. Before formulating our specific research questions, it is important to clarify the applied definition of alliance in the current study. Bordin (1979) defined the alliance as involving consensus on goals and tasks as well as the interpersonal bond. In the current study, we focused on the congruence between therapist and patient therapeutic bond ratings rather than other alliance components. The bond refers specifically to the affective relationship between the patient and therapist as persons and is distinct from the collaborative relationship. Despite recent calls for the investigation of more specific elements of the alliance (e.g., Gelso, 2009; Horvath, 2009), most of the previous studies on congruence of alliance perceptions have been conducted with the rather global working alliance. Therefore, our hypotheses draw mostly on results from these endeavors.

Based on the literature presented above, the following research questions and hypotheses were derived. On the between-dyad level:

1. We expect that therapists' bond ratings will, on average, be significantly lower than their patients' ratings (i.e., reflected in a negative *directional discrepancy*) and that, on average, both perspectives are positively associated with one another (i.e., reflected in a positive *temporal congruence*).
2. We expect that the higher a dyad's temporal congruence of bond ratings, the better their patients' treatment outcomes.
3. A quadratic association between directional discrepancy and treatment outcome is expected. Patients whose therapists rate the therapeutic bond considerably higher or lower than their patients (larger squared discrepancy values) should have worse treatment outcomes than those with similar bond ratings.

On the within-dyad level:

4. Assuming sessions in which patient and therapist bond ratings are in agreement (i.e., patient rating = therapist rating), we expect:
  - a. Higher bond ratings are followed by lower next session symptom ratings (i.e., reflected by a negative  $\alpha$  estimate). For example, sessions in which the patient rating

equals the therapist rating = 0.5 will be followed by lower symptoms ratings than sessions in which the dyad's rating = 1.5.

- b. This association is characterized by a significant quadratic relationship (i.e., reflected in a negative  $a_2$  estimate). As such, agreement on higher bond levels and agreement on lower bond levels is expected to be better (i.e., followed by lower next session symptom scores) than agreement on average bond levels.
5. Assuming sessions in which patient and therapist bond ratings are in disagreement (i.e., patient rating = therapist rating $^*($ -1)), we expect:
- a. Patients' bond ratings will be more important for symptom change than therapists' bond ratings. Consequently, if patients' and therapists' bond ratings are in "disagreement," it should lower next session symptoms are expected if therapists report higher bond ratings than their patients' compared with if patients report higher bond ratings than their therapists (i.e., reflected in a positive  $a_3$  estimate).
  - b. Lower levels of bond disagreement in one session should be associated with lower symptom scores in the subsequent session (i.e., reflected in a positive  $a_4$  estimate). Specifically, if therapists' and patients' bond ratings are too different in one session, irrespective of whether patients or therapists rate the bond more positively than the other, next session symptoms should be lower than if they agree on an average bond level. This hypothesis addresses the test of whether bond agreement has a direct effect on symptom impairment.

## Method

### Participants and Treatment

Data collection was part of the routine outcome monitoring system implemented at the Trier University Outpatient Clinic in Southwest Germany. All patients who started treatment between 2009 and 2015 were considered eligible for this study, as long as they provided at least five HSCL-11 scores (to ensure sufficient within-patient measurement points) and completed the BSI at the beginning and end of treatment. Thus, the analyses were based on a sample of 580 patients treated by 92 therapists. Patients were 36.35 years old on average ( $SD = 12.65$ , range = 15–74) and the majority was female (65.1%). On average, patients received 36 sessions ( $SD = 17.07$ , range: 5–105). Psychological diagnostics were based on the *Structural Clinical Interview for Axis I DSM-IV Disorders* (SCID-I; First, Spitzer, Gibbon, & Williams, 1995), which was conducted before the beginning of therapy by independent clinicians who were trained in SCID administration. Most patients were diagnosed with affective disorders (52.7%) followed by anxiety disorders (18.2%). Other diagnoses were adjustment disorder (9.2%), posttraumatic stress disorder (PTSD; 4.7%), somatoform disorders (4.2%), eating disorders (2.9%), and others (8.1%).

The CBT program consisted of psycho-education on the respective disorder, relaxation training RT, cognitive restructuring (CR) and in sensu as well as in vivo situational exposure for patients with behavioral avoidance. Furthermore, therapists were trained in an integrative CBT approach with interpersonal and emotion-focused elements (Castonguay, Eubanks, Goldfried, Muran, & Lutz, 2015; Grawe, 2004). All therapists were familiar with different published disorder specific CBT manuals, but did individually adapt their approach depending on patients' characteristics. All therapists participated in a 3- to 5-year postgraduate training program with a CBT focus. Therapists were supervised every fourth session by a senior therapist and were supported by a feedback system monitoring patient outcomes on a session-by-session basis. Psychometric feedback was provided to therapists after each session.

### Instruments and Data Collection

**Brief Symptom Inventory (BSI).** Symptom severity was measured using the BSI (Franke, 2000; German translation of Derogatis, 1975), which is a 53-item self-report inventory inquiring about physical and psychological symptoms within the last week. It is the brief form of the Derogatis' Symptom Checklist-90 Revised (SCL-90-R; Derogatis, 1992), which assesses nine subscales with the following dimensions: somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychotism. Item response takes place on a 5-point Likert scale ranging from 0 (*not at all*) to 4 (*extremely*). As the primary outcome measure, a Global Severity Index (GSI) was computed by averaging all BSI items to assess symptom severity both pre- and post-treatment. Psychometric properties for this index can be regarded as excellent ( $\alpha_{pre} = .96$ ;  $\alpha_{post} = .97$ ).

**Hopkins Symptom Checklist—Short Form (HSCL-11).** The HSCL-11 (Lutz, Tholen, Schürch, & Berking, 2006) is an 11-item self-report inventory for the assessment of symptomatic distress. It was developed based on the HSCL-25 (Coyne et al., 1987), which is a brief version of the Hopkins Symptom Checklist-90 (Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974). In the present study, the HSCL-11 was administered at the beginning of each session. The items are responded to on a 4-point Likert scale ranging from 1 (*not at all*) to 4 (*extremely*). The mean of the 11 items represents the patient's level of global symptomatic distress for the preceding week. It is highly correlated with the GSI ( $r = .91$ ) and also has high internal consistency ( $\alpha = .92$ ; Lutz et al., 2006).

**Bern Postsession Reports (BPSR-P/T).** The BPSR (Flückiger, Regli, Zwahlen, Hostettler, & Caspar, 2010) was designed to analyze the process of change as reported by patients (or their therapists) immediately after each session. To assess congruence in bond ratings we chose to use only the two items of the patient and therapist version of the global alliance subscale, which comprise the exact same questions: "The relationship with my therapist [patient] felt comfortable today," "My therapist [patient] and I are getting along well." These items were answered on 7-point Likert scales ranging from -3 (*not at all*) to 3 (*yes, exactly*). Both versions had excellent internal consistencies in the present sample (for patients,  $\alpha$  ranged from .74 to .91 in the first 20 sessions; for therapists,  $\alpha$  ranged from .80 to .89). The BPSR-P/T has been

validated and used in several previous studies (e.g., Flückiger, Grosse Holtforth, Znoj, Caspar, & Wampold, 2013; Grosse Holtforth et al., 2014). As the BPSR and its bond subscale is a relatively short measure, it is well suited for session-to-session assessments. However, its brevity may deem the BPSR less reliable than longer and more established alliance measures. Therefore, to be able to validate BPSR bond scores, patients completed the Penn Helping Alliance Questionnaire (HAQ; Alexander & Luborsky, 1986) every fifth session. For the study sample, the correlation between the two items from the BPSR alliance subscale that were used and the relationship satisfaction scale of the HAQ (those items that were part of both the HAQ and HAQ-II) ranged between  $r = .44$  (Session 25) and  $r = .68$  (Session 20) over the first 30 sessions. These associations strengthen the evidence for the validity of the two-item BPSR subscale used in the current investigation as a measure of the therapeutic alliance. However, given the fact that the HAQ relationship satisfaction subscale also measures the two other alliance components proposed by Bordin (agreement on goals and tasks), it is plausible that the correlations range from medium to large only.

### Data Analytic Strategy

Data analyses were divided into two steps: analyses that are concerned with associations on the between-patient level and analyses that are concerned with associations on the within-patient level. Different analytical models were used to examine these two levels, respectively: We used the T&B model (West & Kenny, 2011) to retrieve patient/dyad level indicators of alliance agreement that can be used to compare patients with different agreement levels with regard to treatment outcome (between-patient level). Polynomial regression and response surface analysis (Edwards & Parry, 1993; Shanock, Baran, Gentry, Pattison, & Heggestad, 2010) is used to assess the direct session-to-session association between alliance agreement and subsequent symptoms (within-patient level).

As the data set has a hierarchical structure where session ratings are nested within patients and patients are nested within therapists, individual observations are not independent of one another and the session-by-session ratings of any patient are more similar to each other than those of two randomly chosen patients. Thus, for both the T&B model and the polynomial regression analysis, we used a three-level MLM (Raudenbush & Bryk, 2002), with sessions at Level 1, patients at Level 2, and therapists at Level 3.

### Truth and Bias Model

In accordance with the approach described in Atzil-Slonim et al. (2015), the truth and bias model was applied to retrieve dyad specific estimates of bond agreement. The model's multilevel equation was as follows:

$$TA_{spt} = (\gamma_{000} + u_{00t} + r_{0pt}) + (\gamma_{100} + u_{10t} + r_{1pt}) * PA_{spt} + e_{spt}$$

The therapists' bond rating ( $TA_{spt}$ ) for session  $s$  of patient  $p$ , who was treated by therapist  $t$ , was predicted by the sample average (i.e., fixed) *directional discrepancy* (i.e., the intercept;  $\gamma_{000}$ ) and *temporal congruence* (i.e., the slope;  $\gamma_{100}$ ) multiplied by this patient's bond rating in that session (i.e.,  $PA_{spt}$ ). *Directional discrepancy* and *congruence* effects are allowed to vary between

therapists (i.e., random effects at Level 3;  $u_{00t}$ ,  $u_{10t}$ ), and between patients (i.e., random effects at Level 2;  $r_{0ct}$ ,  $r_{1ct}$ ). The residual term quantifies the session-specific deviation from these effects (i.e., random effect at Level 1;  $e_{spt}$ ). For each patient-therapist dyad, an empirical Bayes estimate of *directional discrepancy* (i.e.,  $u_{00t} + r_{0pt}$ ) and *congruence* (i.e.,  $u_{10t} + r_{1pt}$ ) effects is retrieved and used in subsequent multiple regression analyses, in which their associations with treatment outcome (BSI posttreatment scores) are analyzed.

As West and Kenny (2011, pp. 374–375) suggested, we centered both the outcome (i.e., therapists' reports) and the predictor (i.e., the patients' reports) on the predictor's person-mean (i.e., each patient's mean reports) across all sessions. This approach allowed us to remove broad individual differences when examining within-person fluctuations. Moreover, as the intercept represents the outcome level when the predictor is set to zero (i.e., when the patient's report is on its average level), centering the outcome on the predictor's person-mean allows the intercept estimate to represent the *directional discrepancy*—that is, the average difference between the therapists' reports and their patients' reports. In other words, the degree to which therapists rate the bond higher (in cases of positive intercepts) or lower (in cases of negative ones) than their patients. Given this centering, patients' bond ratings had a mean of  $M = 0$  (average within-dyad  $SD = 0.45$ ). Therapists rated the bond on average  $M = -0.42$  (average within-dyad  $SD = 0.57$ ) points below the respective patient mean.

### Polynomial Regression and Response Surface Analysis

To examine associations on the within-patient level, we used the polynomial regression version of the response surface analysis. RSA allows the examination of the combined effect of two predictor variables on one outcome variable. These effects can be visualized in a three-dimensional space (Edwards & Parry, 1993; Shanock et al., 2010). The model's multilevel equation was as follows:

$$\begin{aligned} HSCL_{s+1pt} = & (\gamma_{000} + u_{00t} + r_{0ct}) + \gamma_{100} * HSCL_{spt} \\ & + \gamma_{200} * PA_{spt} + \gamma_{300} * TA_{spt} \\ & + \gamma_{400} * (PA_{spt} * TA_{spt}) + \gamma_{500} * (PA_{spt})^2 \\ & + \gamma_{600} * (TA_{spt})^2 + e_{spt}; \end{aligned}$$

The symptom rating for session  $s + 1$  of patient  $p$ , who was treated by therapist  $t$  (i.e.,  $HSCL_{s+1pt}$ ) was predicted by this patient's preceding symptom impairment in session  $s$  (i.e.,  $\gamma_{100} * HSCL_{spt}$ ), this patient's bond ratings in session  $s$  (i.e.,  $\gamma_{200} * PA_{spt}$ ), the bond ratings of the therapist  $t$  treating this patient  $c$  (i.e.,  $\gamma_{300} * TA_{spt}$ ), the interaction of these bond ratings (i.e.,  $\gamma_{400} * (PA_{spt} * TA_{spt})$ ), and the squared bond ratings of the patient and the therapist (i.e.,  $\gamma_{500} * (PA_{spt})^2$  and  $\gamma_{600} * (TA_{spt})^2$ , respectively). Additionally, the random terms  $u_{00t}$  and  $r_{0ct}$  control for the nested structure of the data (sessions nested within patients, nested within therapists). As for the T&B model analyses described above, the patient's and therapist's bond ratings were centered around the patient's person-mean score.

In response surface analysis, the coefficients  $\gamma_{200}$  to  $\gamma_{600}$  are used to calculate test values for two slopes and two respective curvatures along the response surface: ( $a_1$ ) The slope along the line of agreement ( $a_1 = \gamma_{200} + \gamma_{300}$ ); this slope reveals if and how bond ratings are associated with next session symptoms if therapist

and patient ratings are in agreement ( $a_2$ ). The curvature along the line of agreement ( $a_2 = \gamma_{400} + \gamma_{500} + \gamma_{600}$ ); this curvature reveals whether the association of bond ratings and next session symptoms is quadratic, that is, if the association given agreement varies depending on the bond ratings or if the association is equal over the range of bond ratings ( $a_3$ ). The slope along the line of disagreement ( $a_3 = \gamma_{200} - \gamma_{300}$ ); this slope reveals whether in the case of disagreement it is worse for next session symptom severity if therapists rate the bond high and patients low or vice versa ( $a_4$ ). The curvature along the line of disagreement ( $a_4 = \gamma_{400} - \gamma_{500} + \gamma_{600}$ ); this curvature reveals whether next session symptoms are lower if disagreement is lower (i.e., agreement is higher).

## Results

### Between-Dyad Associations of Congruence and Discrepancy With Symptoms

Results from the truth and bias model analysis supported Hypothesis 1. On average, therapists rated the therapeutic bond lower than their patients by  $\gamma_{000} = -.42$  ( $SE = .03$ ,  $t = -15.89$ ,  $p < .001$ ) points. Additionally, the average congruence between patients' and therapists' bond ratings was  $\gamma_{100} = .63$  ( $SE = .02$ ,  $t = 39.66$ ,  $p < .001$ ).

To test our Hypotheses 2 and 3, the individual dyad estimates for congruence and discrepancy were retrieved and their associations with treatment outcome tested in a multilevel regression analysis (patients nested within therapists).<sup>1</sup> Supporting our Hypotheses 2 and 3, Table 1 shows that the congruence ( $\gamma_{30}$ ) as well as the squared discrepancy ( $\gamma_{40}$ ) had a significant association with  $GSI_{post}$  scores, while adjusting for  $GSI_{pre}$  scores. The higher the congruence between therapists' and patients' bond ratings over the course of treatment, the lower patients' symptom distress at post treatment. Discrepancy shows a quadratic association with treatment outcome. The first derivative of the quadratic discrepancy function ( $f'(discrepancy) = 2*0.32*discrepancy - 0.09 = >$  minimum at discrepancy = 0.14) indicates, that the lowest posttreatment symptom severity score was expected if therapists' bond ratings were on average 0.28 points lower than patients' bond ratings (i.e., 0.14 points above the average discrepancy score of  $-0.42$ ). Deviations above and below that optimal level were associated with worse treatment outcome.

Table 1  
Multilevel Regression Results for the Prediction of Treatment Outcome by Congruence and Bias

Effect	Estimate	SE	t	p
Intercept $\gamma_{00}$	.583	.025	23.672	<.001***
$GSI_{pre}$ $\gamma_{10}$	.416	.031	13.501	<.001***
Discrepancy $\gamma_{20}$	-.089	.070	-1.276	.203
Congruence $\gamma_{30}$	-.29	.108	-2.677	.008**
Discrepancy <sup>2</sup> $\gamma_{40}$	.318	.111	2.868	.004**

Note.  $GSI$  = Global Severity Index.

\*\*  $p < .01$ . \*\*\*  $p < .001$ .

### Within-Dyad Associations of Agreement/Disagreement With Symptoms

Table 2 shows the results of the polynomial regression analysis. Patients' ( $\gamma_{200}$ ) as well as therapists' ( $\gamma_{300}$ ) bond ratings showed a significant negative association with next session symptom impairment, when controlling for symptoms in the current session. Additionally, we found a significant quadratic association between patients' bond ratings and symptoms ( $\gamma_{600}$ ).

The four values ( $a1$ – $a4$ ) that result when subjecting these coefficients to response surface analysis are presented in Table 3 and the complete response surface is displayed in Figure 1. The significant slope along the line of agreement ( $a1$ ) supports our Hypothesis 4b and indicates that if patients' and therapists' bond ratings are in agreement, higher bond scores are associated with lower symptom impairment in subsequent sessions, when controlling for symptoms in the current session (see Figure 2). Along the line of disagreement, we found a significant negative linear relationship ( $a3$ ), as well as a significant curvature ( $a4$ ), which contradicts our Hypotheses 5a and 5b (see Figure 3). Following that, if patients and their therapists perceive the bond differently, next session symptoms are less severe if the therapist rates the bond as stronger than if the patient rates the bond as stronger. Given the significantly negative  $a4$  term, subsequent symptoms are less severe if patients' and therapists' ratings are in disagreement, that is one of the two perceives the bond as strong while the other perceives it as weak, than if both parties agree on a person-specific average bond level.<sup>2</sup>

## Discussion

The present study investigated the associations of congruence of patients' and therapists' bond ratings with patients' symptom distress on the within- and between-dyad level using session-to-session data. Previous research has produced mixed results using either limited statistical approaches, or only one or few alliance measurements over the course of the treatment. We used two sophisticated methods and intensive longitudinal measurements to overcome the aforementioned limitations of prior research: The truth and bias model was used to generate dyad specific estimates of temporal congruence (i.e., mutual fluctuations) and directional discrepancy (i.e., average differences) of bond ratings over the course of treatment. These estimates were used to test whether dyads with more congruent bond ratings and lower discrepancy had more positive treatment outcomes. Polynomial regression and response surface analyses were utilized to investigate the associations between bond agreement/disagreement and next session

<sup>1</sup> A random intercept model was preferred over more complex models, as inclusion of further random terms resulted in nonconvergence.

<sup>2</sup> In an attempt to examine some of the potential reasons for the mixed results regarding the congruence-outcome association, we conducted an exploratory analysis to determine whether treatment length impacts the results of our analyses. For the between-dyad associations presented in Table 1 we, therefore, included interactions with treatment length to the model predicting  $GSI_{post}$  scores. Neither temporal congruence nor squared directional discrepancy showed a significant interaction with length of treatment (i.e., number of sessions attended). Similarly, for the within-dyad associations, we tested interactions between the predictors included in the model displayed in Table 2 and treatment length. Again, none of the predictors showed a significant interaction effect with treatment length.

Table 2  
*Polynomial Regression Model for Patient and Therapist Alliance and Patient-Rated Symptom Severity From Session to Session*

Effect	Estimate	SE	<i>t</i>	<i>p</i>
Intercept $\gamma_{000}$	1.778	.0194	91.701	<.001***
HSCL- $\gamma_{100}$	.483	.006	77.779	<.001***
Therapist bond $\gamma_{200}$	-.064	.007	-9.705	<.001***
Patient bond $\gamma_{300}$	-.026	.006	-4.087	<.001***
Therapist $\times$ Patient bond $\gamma_{400}$	.0001	.002	.080	.936
Therapist bond <sup>2</sup> $\gamma_{500}$	-.003	.003	-.825	.409
Patient bond <sup>2</sup> $\gamma_{600}$	-.011	.004	-2.368	.018*

Note. HSCL = Hopkins Symptom Checklist.

\*  $p < .05$ . \*\*\*  $p < .001$ .

symptom impairment (while controlling for current symptom impairment) on the within-dyad session level.

Replicating previous results (e.g., Atzil-Slonim et al., 2015; Compare et al., 2016; Fitzpatrick et al., 2005; Tryon et al., 2007; Zilcha-Mano, Snyder, & Silberschatz, 2016) and supporting our first hypothesis, we found that, on average, therapists rate the bond lower than their patients (negative directional bias), but both parties similarly perceive changes in the bond over the course of the treatment (positive congruence).

In support of our second hypothesis, we found a significant between-dyad association between congruence of bond ratings and treatment outcome. That is, the higher a dyad's bond congruence, the lower their patients' symptoms at the end of treatment (while controlling for symptoms at the start of treatment). This finding underlines the importance of therapists' perceiving changes in the therapeutic bond similarly to their patients. Similar results have been reported by the only other study that investigated the association between mutual fluctuations in session-to-session alliance ratings and treatment outcome in adult individual psychotherapy (Rozmarin et al., 2008). However, because Compare et al. (2016) did not find this association in a sample of patients treated with emotion-focused group therapy, this finding may be specific for dyads from individual psychotherapy. The relationship build between patients and therapists in a group setting may plausibly be structurally different from the more intimate relationship in individual settings (cf. Piper & Ogrodniczuk, 2010). Of course, there are several other potential explanations for these divergent findings like differences in alliance measures, treatment approaches, and diagnostic groups. More specific research is needed to unravel the association between temporal congruence and outcome for different settings, orientations, and disorders.

In support of our third hypothesis, a significant association was found between squared directional bias and posttreatment symptoms. Lowest GSI<sub>post</sub> scores were present for therapists' bond ratings that were moderately lower than those of their patients. Therapists' bond ratings higher or lower than 0.28 points below their patients' ratings were associated with more severe symptom scores after treatment. For a dyad to achieve an average directional discrepancy scores of -0.28 (i.e., therapists rate the bond on average 0.28 points lower than their patients), which is estimated as optimal in the current study, both parties need to provide rather overlapping bond ratings over the course of the treatment. For example, in a treatment with an average length (i.e., 36 sessions in the current sample), a discrepancy of -0.28 would be achieved if

therapists provided a bond rating one point below the patient's rating in 10 sessions and identical ratings in 26 sessions. Of course, this is only one example and there are ample different possibilities of how dyads can eventually end up with a discrepancy of -0.28. However, what becomes apparent within this numerical example is that both have to report fairly similar ratings over the course of the treatment. These results provide additional support for the importance of a shared perception of the therapeutic bond, but also the importance of therapists being sensitive to detect problems in the bond that patients might not be aware of (e.g., Fjermstad et al., 2016; Rozmarin et al., 2008; Zilcha-Mano, Snyder, & Silberschatz, 2016).

More important, these between-dyad bond congruence-outcome associations could be the result of many unmeasured potentially confounding influences. This is a typical problem of psychotherapy process-outcome research in which a random assignment of patients or dyads to different groups is often not feasible (e.g., Falkenström et al., 2017). For example, therapists' attachment could have an influence on their alliance ratings (Kivlighan & Marmarosh, 2016). Thus, our next step was to explore these associations on a within-dyad level at which stable characteristics of the dyad can be ruled out as alternative explanations of the results.

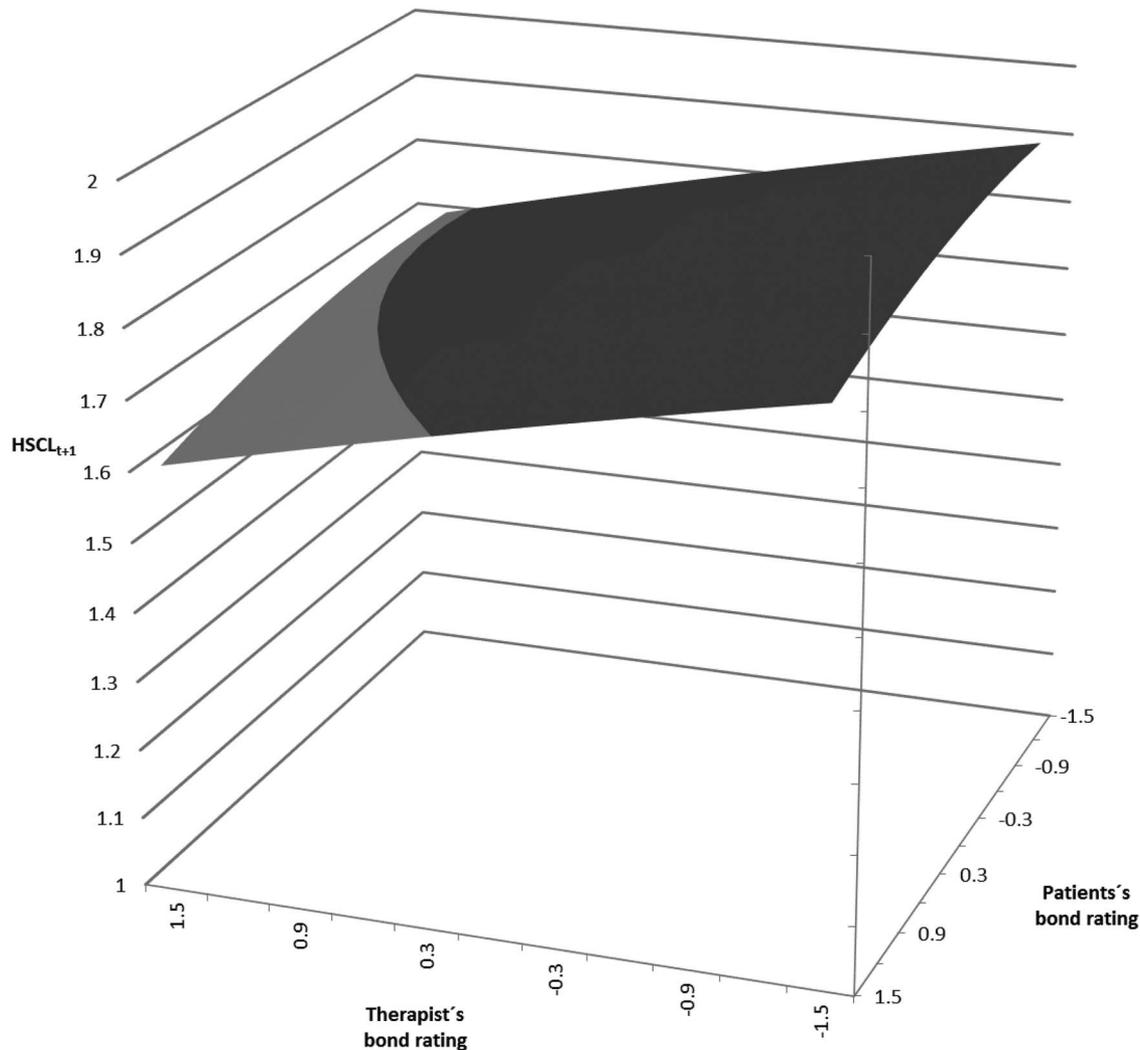
On the within-dyad level, results of the response surface analysis supported our Hypothesis 4a—namely, that if patient and therapist bond ratings are in agreement, higher bond ratings are associated with lower next session symptoms, when controlling for current symptoms. These results replicate Zilcha-Mano, Snyder, and Silberschatz's (2016) findings, although these authors used a larger lag between alliance and symptom ratings (1 month instead of one session) and did not control for symptom impairment in the session, in which the alliance ratings were assessed. However, we could not replicate Zilcha-Mano, Snyder, and Silberschatz's (2016a) finding that this association has a significant curvature (Hypothesis 4b). This curvature would imply that patients profit more from sessions in which therapists and patients agree on a low alliance or on a high alliance than if they agree on an average alliance quality. The authors explained this finding with the respective implications these different forms of alliance agreement may have. Agreement on a high alliance level might indicate an optimal environment for therapeutic work, whereas agreement on a low alliance level might indicate the need to specifically aim at resolving problems in the alliance (cf. Safran & Muran, 1996). Agreement on a moderate alliance level may neither indicate a

Table 3  
*Response Surface for Patient and Therapist Alliance Ratings and Next Session Symptoms Controlling for Current Symptoms*

Effect	Coefficient	SE	<i>t</i> (20,835)	<i>p</i>
a1: Slope along $x = y$	-.09	.01	-12.203	<.001***
a2: Curvature along $x = y$	-.01	.00	-1.803	.054
a3: Slope along $x = -y$	-.04	.01	-3.653	<.001**
a4: Curvature along $x = -y$	-.01	.00	-2.868	<.01**

Note. Slope along = association of patient and therapist bond ratings (given agreement:  $x = y$  and disagreement:  $x = -y$ ); curvature along = quadratic trend of the line of agreement and disagreement.

\*\*  $p < .01$ . \*\*\*  $p < .001$ .



*Figure 1.* Response surface displaying the within-patient association of patient and therapist bond ratings and next session symptoms while controlling for current symptoms ( $HSCL_t$ ). The  $x$ -axis is the patient's bond, the  $y$ -axis is the therapist's bond, and the  $z$ -axis is the patient's symptom scores in the following session adjusted for symptoms in the current session.

good environment for symptom-oriented work nor the necessity to resolve alliance ruptures and, thus, may leave therapists guessing what the right next steps are (cf. Zilcha-Mano, Snyder, & Silberschatz, 2016). Besides the different timing of assessments over the course of treatment (session-to-session vs. every fifth session), another explanation for the contrasting findings in the present study may again be the different treatment approaches. CBT therapists in the present sample may be less focused on the resolution of alliance ruptures than the cognitive-relational psychodynamic therapists in the Zilcha-Mano, Snyder, and Silberschatz (2016) study. Future research needs to investigate whether training therapists' abilities to negotiate alliance ruptures results in patients profiting more from sessions in which both agree on a low alliance quality (e.g., Muran, Safran, & Eubanks-Carter, 2010).

In contrast to our Hypothesis 5a, we found that if patient and therapist bond ratings were in disagreement, it was worse if ther-

apists rated the quality of the bond low and patients high than vice versa. That is, given disagreement in bond ratings (patients bond = therapists bond<sup>\*</sup>(-1)), next session symptoms were more severe if the patient perceived the bond as stronger than if the therapist perceived the bond as stronger.

This surprising finding is difficult to understand clinically and inconsistent with many results of previous research. For example, there is a broad body of literature in which the actor-partner-interdependence model (APIM; Kashy & Kenny, 2000) is applied to investigate the association of patient-rated alliance (*actor effect*) and therapist rated alliance (*partner effect*) with patient-rated symptoms (e.g., Gelso et al., 2012; Kivlighan, Gelso, Ain, Hummel, & Markin, 2015; Kivlighan et al., 2016; Kivlighan, Marmarosh, & Hilsenroth, 2014; Markin, Kivlighan, Gelso, Hummel, & Spiegel, 2014; Zilcha-Mano et al., 2016). In most of these aforementioned studies, actor effects were stronger than partner effects.

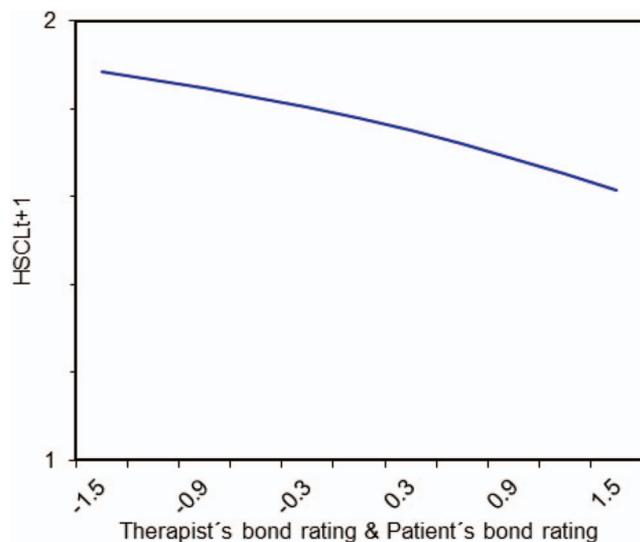


Figure 2. Within-dyad association of patient and therapist bond ratings (given agreement) and next session symptom ratings while controlling for current symptoms ( $HSCL_t$ ). See the online article for the color version of this figure.

Because, in the present study, symptoms were measured by patient self-reports, patient bond-symptom associations can be regarded as an actor effect and therapist bond-symptom associations can be regarded as a partner effect. Additionally, our finding is in contrast to the notion of the “single rater bias” assuming higher associations between ratings provided by the same source than those of different sources because of shared method variation (e.g., [Hepner, Wampold, & Kivlighan, 2008](#); [Orth, 2013](#)).

Given the considerable amount of competing findings, our result regarding the priority of therapists’ over patients’ bond ratings should be interpreted with caution. Methodological or sample specific reasons might hamper the generalizability of our finding. For example, it is important to keep in mind that sessions in which therapists rate the bond better than their patients are rather seldom. In this sample, therapists’ bond ratings were higher than their patients’ ratings in only 10.6% of the sessions. As such, firm conclusions about the associations of this scenario with symptoms should not be drawn based on this one study. In future research, it could be especially helpful to include observer-based assessments with regard to bond ratings in sessions of disagreement, allowing for a more fine grained analysis of what happens in these sessions. In sessions in which therapists perceive the bond as strong, they may rather engage in coping-oriented interventions, which could facilitate symptom change, although patients perceive a weak affective relationship with their therapist. Another possibility is that in sessions in which the therapist rates the bond lower than the patient, they are accurately perceiving a problem in the bond, while the patient rates the bond high because he or she is overly deferential, appealing, or in denial about the problem in the bond. This last explanation describes phenomena that are known in the literature as withdrawal ruptures ([Harper, 1989](#); [Safran & Muran, 1996](#)). Although rupture-repair episodes have been shown to be positively associated with treatment outcome in the long run, resolution processes may need several sessions after the rupture to

develop (e.g., [Larsson, Falkenström, Andersson, & Holmqvist, 2016](#); [Safran, Muran, & Eubanks-Carter, 2011](#); [Strauss et al., 2006](#)). As such, these ruptures might be associated with higher symptoms directly after the impasse, when resolution has not yet begun or is not completed.

In contrast to our Hypothesis 5b, we did find that lower levels of disagreement in one session were related to higher symptom scores in the next session. That is, sessions were rather more productive if therapists perceived the bond as strong and patients as weak than if both agreed on an average bond quality. As this effect leveled out for lower therapist bond ratings (i.e., higher patient ratings in the case of disagreement), this association was especially pronounced for sessions in which therapists perceived a strong bond. If not so much the perception of the bond by therapist and patient, but agreement on the quality of the bond were key for the promotion of therapeutic change, we would have expected an inverse U-shaped line of disagreement. That is, sessions in which patients and therapists agreed on an average bond level should have been more productive than sessions in which they perceived the quality of the bond differently. As this is the first investigation of this association on a session-by-session within-dyad level, we can only compare our results with research from between-dyad comparisons. In accordance with our results, neither [Marmarosh and Kivlighan \(2012\)](#) nor [Zilcha-Mano, Snyder, and Silberschatz \(2016\)](#) found a U-shaped association in the line of disagreement in patient and therapist perceptions of the working alliance and session smoothness or symptoms. On the other hand, [Fjermstad et al. \(2016\)](#) found that therapist-youth dyads with higher levels of agreement on change in the alliance from the third to the seventh session of CBT anxiety treatments produced better long-term outcomes. However, these effects were only present at 1-year follow-up and not immediately after the treatment. Given this evidence supporting the idea that the effects of alliance congruence

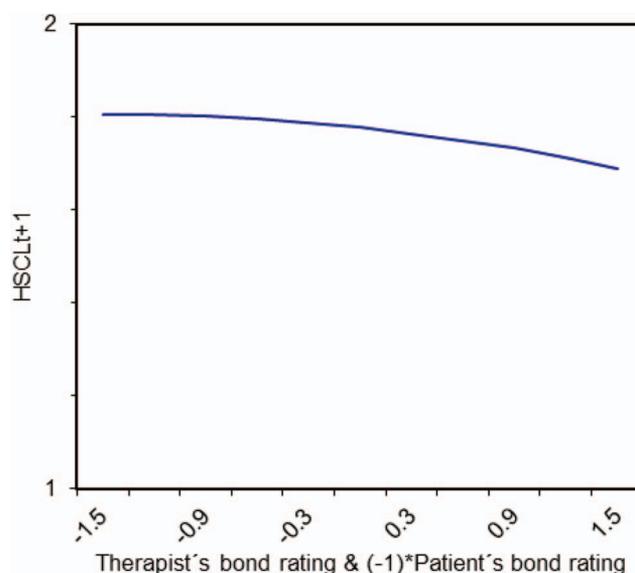


Figure 3. Within-dyad association of patient and therapist bond ratings (given disagreement) and next session symptom ratings while controlling for current symptoms ( $HSCL_t$ ). See the online article for the color version of this figure.

may take a longer time to manifest, it does not come as a surprise that we did not find an immediate session-to-session association between bond congruence and symptoms.

Overall our results suggest that therapists provide more immediately helpful interventions when they perceive the bond as strong compared with when they perceive the bond as weak. If patients agree with their therapists on this high bond quality this seems to further enhance the productivity of the session. However, agreement in itself seems not to result in immediate symptom improvements. Rather the overall amount of congruence and discrepancy in patient and therapist bond ratings over the entire course of treatment seems to be important for a dyad to achieve therapeutic change. These associations seem to be independent of the length of the treatment.

### Limitations, Future Directions, and Summary

One limitation of our study is that all therapists received routine feedback on their patients' symptom and bond ratings after each session. Consequently, their bond ratings in subsequent sessions may be biased by their knowledge of how their patients perceived the bond in previous sessions. However, it is important to notice the timing of the feedback and the assessments. Therapists only received the feedback on their patients' bond ratings after they had rated the alliance for a respective session, so that they were unaware of the patients' ratings at the time they provided their bond ratings. Most of the recent studies investigating alliance agreement-outcome associations collected data in settings, in which therapists received feedback on their patients' symptom change and/or alliance ratings (Atzil-Slonim et al., 2015; Zilcha-Mano, Snyder, & Silberschatz, 2016). Future studies, which also include groups of therapists who are not provided with feedback on their patients' alliance ratings are needed to test whether this feedback in-/decreases congruence and bias as well as their effects on symptom impairment.

Furthermore, one can argue that therapists may not be guided by how they perceive the quality of the bond, but how they think their patients perceive it. Most alliance instruments ask therapists about their impression of the alliance and not about their view of how their patients perceive it. Consequently, only few studies thus far have investigated the congruence of patient alliance ratings and therapist judgments of these ratings (Hartmann et al., 2015; Ormhaug, Shirk, & Wentzel-Larsen, 2015). However, this distinction assumes that therapists have two more or less independent views on the alliance. There is some evidence that casts doubt about therapists being able to make this distinction (Creed & Kendall, 2005). In a study on alliance-building behavior of CBT therapists treating 56 youth with anxiety, Creed and Kendall (2005) found a correlation of  $r = .95$  between how therapists thought the child viewed the alliance and how therapists themselves viewed the alliance. This high correlation lends support to the suggestion that therapists' perception of the therapeutic bond may well represent how therapists think their patients perceive the bond. Thus, it seems plausible that therapists use their own perception of the bond as a tool to guide them in clinical decision-making.

Another potential limitation of the current study regards the use of the bond component of the alliance. Future studies may benefit from examining the association between congruence of therapist and patient ratings of other facets of the therapeutic alliance

(namely, tasks and goals) and treatment outcome. More important, instead of the global alliance subscale of the BPSR, which was used in previous studies (e.g., Atzil-Slonim et al., 2015), we only used the two respective overlapping items of the patient and therapist subscales.

Moreover, it should be recognized that symptoms are solely assessed from the patients' perspective. While self-report measures reflect an important source for the assessment of treatment progress, therapists' or independent observers' ratings may also reveal important aspects of the change process, which do not always concur with patients' ratings (e.g., Ogles, 2013). Consequently, the results of the present study should not be generalized without caution to circumstances in which other means of outcome assessment are used.

It should further be noted that the repeated assessment of bond and symptoms itself may have effects on the use of these self-reports beyond changes in the very construct, which is to be assessed. Recollection biases could induce a lack of variation in the measures and thus an artificial stability in the constructs under scrutiny. However, we are not aware of a suitable statistical method that on the one hand enables researchers to take advantage of the rich information repeated measurement designs provide and on the other hand quantify or control for these statistical artifacts.

Finally, it is important to note that the associations found in this study might be different depending on the specific composition of the patient sample. Especially diagnostic categories (Axis I as well as Axis II) might be potential moderators of the reported associations. Future research that more specifically focuses on these aspects is needed to unravel disorder specific relationships.

Notwithstanding the limitations noted above, the current study is the first to use two contemporary methods for the study of bond agreement, namely the T&B model and RSA, in a large dataset of patients receiving individual cognitive-behavioral therapy to investigate its associations with treatment outcome on the within- and between-dyad level. The results of this study encourage the incorporation of patients' session-by-session bond ratings into routine monitoring and feedback systems (e.g., Lambert, 2007). In the past, these feedback systems have been primarily concerned with outcome measures alone and research has provided evidence for the importance of therapists' accurately tracking patients' functioning over the course of the treatment (e.g., Bar-Kalifa et al., 2016; Lutz, DeJong, & Rubel, 2015). Additionally, these results can inform the development of decision rules based on alliance feedback (e.g., Lutz, Zimmermann, Müller, Deisenhofer, & Rubel, 2017). For example, with regard to the absolute difference between patient and therapist bond ratings, it would be possible to define a kind of "optimal zone" in which both perspectives are in an optimal relation. Future research must test the additional value of such bond feedback in randomized controlled trials and naturalistic studies.

### References

- Alexander, L. B., & Luborsky, L. (1986). The Penn Helping Alliance Scales. In L. S. Greenberg & W. M. Pinsof (Eds.), *The psychotherapeutic Process: A research handbook* (pp. 325–366). New York, NY: Guilford Press.
- Atzil-Slonim, D., Bar-Kalifa, E., Rafaeli, E., Lutz, W., Rubel, J., Schiefele, A.-K., & Peri, T. (2015). Therapeutic bond judgments: Congruence and

- incongruence. *Journal of Consulting and Clinical Psychology*, 83, 773–784. <http://dx.doi.org/10.1037/ccp0000015>
- 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>
- Bachelor, A., & Salamé, R. (2000). Participants' perceptions of dimensions of the therapeutic alliance over the course of therapy. *Journal of Psychotherapy Practice and Research*, 9, 39–53.
- Bar-Kalifa, E., Atzil-Slonim, D., Rafaeli, E., Peri, T., Rubel, J., & Lutz, W. (2016). Therapist-client agreement in assessments of clients' functioning. *Journal of Consulting and Clinical Psychology*, 84, 1127–1134. <http://dx.doi.org/10.1037/ccp0000157>
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice*, 16, 252–260. <http://dx.doi.org/10.1037/h0085885>
- Castonguay, L. G., Constantion, M. J., McAleavey, A. A., & Goldfried, M. R. (2010). The therapeutic alliance in cognitive-behavioral therapy. In J. Muran & J. Barber (Eds.), *The therapeutic alliance: An evidence-based guide to practice* (pp. 150–171). New York, NY: Guilford Press.
- Castonguay, L. G., Eubanks, C. F., Goldfried, M. R., Muran, J. C., & Lutz, W. (2015). Research on psychotherapy integration: Building on the past, looking to the future. *Psychotherapy Research*, 25, 365–382. <http://dx.doi.org/10.1080/10503307.2015.1014010>
- Compare, A., Tasca, G. A., Lo Coco, G., & Kivlighan, D. M. (2016). Congruence of group therapist and group member alliance judgments in emotionally focused group therapy for binge eating disorder. *Psychotherapy: Theory, Research & Practice*, 53, 163–173. Advance online publication. <http://dx.doi.org/10.1037/pst0000042>
- Coyen, J. C., Kessler, R. C., Tal, M., Turnbull, J., Wortman, C. B., & Greden, J. F. (1987). Living with a depressed person. *Journal of Consulting and Clinical Psychology*, 55, 347–352. <http://dx.doi.org/10.1037/0022-006X.55.3.347>
- Creed, T. A., & Kendall, P. C. (2005). Therapist alliance-building behavior within a cognitive-behavioral treatment for anxiety in youth. *Journal of Consulting and Clinical Psychology*, 73, 498–505. <http://dx.doi.org/10.1037/0022-006X.73.3.498>
- Crits-Christoph, P., Gibbons, M. B. C., & Mukherjee, D. (2013). Psychotherapy process-outcome research. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 298–340). New York, NY: Wiley.
- 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. (1975). *Brief Symptom Inventory*. Baltimore, MD: Clinical Psychometric Research.
- Derogatis, L. R. (1992). *SCL-90-R: Administration, scoring & procedures manual-II for the (revised) version and other instruments of the psychopathology rating scale series* (2nd ed.). Towson, MD: Clinical Psychometric Research.
- Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins symptom checklist (HSCL): A self-report symptom inventory. *Behavioral Science*, 19(1), 1e15. <http://dx.doi.org/10.1002/bs.3830190102>
- Edwards, J. R., & Parry, M. E. (1993). On the use of polynomial regression equations as an alternative to difference scores in organizational research. *Academy of Management Journal*, 36, 1577–1613. <http://dx.doi.org/10.2307/256822>
- Falkenström, F., Ekeblad, A., & Holmqvist, R. (2016). Improvement of the working alliance in one treatment session predicts improvement of depressive symptoms by the next session. *Journal of Consulting and Clinical Psychology*, 84, 738–751. <http://dx.doi.org/10.1037/ccp0000119>
- Falkenström, F., Finkel, S., Sandell, R., Rubel, J. A., & Holmqvist, R. (2017). Dynamic models of individual change in psychotherapy process research. *Journal of Consulting and Clinical Psychology*, 85, 537–549. <http://dx.doi.org/10.1037/ccp0000203>
- 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>
- Fenton, L. R., Cecero, J. J., Nich, C., Frankforter, T. L., & Carroll, K. M. (2001). Perspective is everything: The predictive validity of six working alliance instruments. *The Journal of Psychotherapy Practice and Research*, 10, 262–268.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1995). *Structured clinical interview for axis I DSM-IV disorders-patient Ed. (SCID-I/P)*. New York, NY: Biometrics Research Department, NY State Psychiatric Institute.
- Fitzpatrick, M. R., Iwakabe, S., & Stalikas, A. (2005). Perspective divergence in the alliance. *Psychotherapy Research*, 15, 69–80. <http://dx.doi.org/10.1080/10503300512331327056>
- Fjermestad, K. W., Lerner, M. D., McLeod, B. D., Wergeland, G. J. H., Heiervang, E. R., Silverman, W. K., . . . Haugland, B. S. (2016). Therapist-youth agreement on alliance change predicts long-term outcome in CBT for anxiety disorders. *Journal of Child Psychology and Psychiatry*, 57, 625–632. <http://dx.doi.org/10.1111/jcpp.12485>
- 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. <http://dx.doi.org/10.1037/a0025749>
- Flückiger, C., Grosse Holtforth, M., Znoj, H. J., Caspar, F., & Wampold, B. E. (2013). Is the relation between early post-session reports and treatment outcome an epiphenomenon of intake distress and early response? A multi-predictor analysis in outpatient psychotherapy. *Psychotherapy Research*, 23(1), 1e13. <http://dx.doi.org/10.1080/10503307.2012.693773>
- Flückiger, C., Regli, D., Zwahlen, D., Hostettler, S., & Caspar, F. (2010). The Bern Post Session Report 2000, patient and therapist versions: Measuring psychotherapeutic processes. *Zeitschrift für Klinische Psychologie & Psychotherapie*, 39, 71–79.
- Franke, G. (2000). *BSI. Brief symptom inventory: Deutsche version. Manual*. Göttingen, Germany: Beltz.
- Gelso, C. J. (2009). The time has come: The real relationship in psychotherapy research. *Psychotherapy Research*, 19, 278–282. <http://dx.doi.org/10.1080/10503300902777155>
- 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>
- Grawe, K. (2004). *Psychological therapy*. Seattle, WA: Hogrefe.
- Grosse Holtforth, M., Altenstein, D., Krieger, T., Flückiger, C., Wright, A. G., & Caspar, F. (2014). Interpersonal differentiation within depression diagnosis: Relating interpersonal subgroups to symptom load and the quality of the early therapeutic alliance. *Psychotherapy Research*, 24, 429–441. <http://dx.doi.org/10.1080/10503307.2013.829253>
- Harper, H. (1989). *Coding guide II: Identification of withdrawal challenges in exploratory therapy*. Sheffield, England: University of Sheffield.
- Hartmann, A., Joos, A., Orlinsky, D. E., & Zeeck, A. (2015). Accuracy of therapist perceptions of patients' alliance: Exploring the divergence. *Psychotherapy Research*, 25, 408–419. <http://dx.doi.org/10.1080/10503307.2014.927601>
- Heppner, P. P., Wampold, B. E., & Kivlighan, D. M., Jr. (2008). *Research design in counseling* (3rd ed.). Belmont, CA: Thomson Brooks/Cole.
- Hoffart, A. (2014). The need to study within-person processes in psychotherapy: A discussion of alliance studies. *SOJ Psychology*, 1, 1–5.

- Hoffart, A., Øktedalen, 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. O. (2009). How real is the “real relationship”? *Psychotherapy Research, 19*, 273–277. <http://dx.doi.org/10.1080/10503300802592506>
- Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy: Theory, Research & Practice, 48*, 9–16. <http://dx.doi.org/10.1037/a0022186>
- 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., 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., Hill, C. E., Gelso, C. J., & Baumann, E. (2016). Working alliance, real relationship, session quality, and client improvement in psychodynamic psychotherapy: A longitudinal actor partner interdependence model. *Journal of Counseling Psychology, 63*, 149–161. <http://dx.doi.org/10.1037/cou0000134>
- Kivlighan, D. M., Jr., & Marmarosh, C. L. (2016). Counselors’ attachment anxiety and avoidance and the congruence in clients’ and therapists’ working alliance ratings. *Psychotherapy Research*. Advance online publication. <http://dx.doi.org/10.1080/10503307.2016.1198875>
- 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. <http://dx.doi.org/10.1037/a0034939>
- Lambert, M. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research, 17*, 1–14. <http://dx.doi.org/10.1080/10503300601032506>
- Langhoff, C., Baer, T., Zubraegel, D., & Linden, M. (2008). Therapist–patient alliance, patient–therapist alliance, mutual therapeutic alliance, therapist–patient concordance, and outcome of CBT in GAD. *Journal of Cognitive Psychotherapy, 22*, 68–79. <http://dx.doi.org/10.1891/0889.8391.22.1.68>
- Larsson, M. H., Falkenström, F., Andersson, G., & Holmqvist, R. (2016). Alliance ruptures and repairs in psychotherapy in primary care. *Psychotherapy Research, 28*, 123–136.
- Lutz, W., De Jong, K., & Rubel, J. (2015). Patient-focused and feedback research in psychotherapy: Where are we and where do we want to go? *Psychotherapy Research, 25*, 625–632. <http://dx.doi.org/10.1080/10503307.2015.1079661>
- Lutz, W., Tholen, S., Schürch, E., & Berking, M. (2006). Die entwicklung, validierung und reliabilität von kurzformen gängiger psychometrischer instrumente zur evaluation des therapeutischen fortschritts in psychotherapie und psychiatrie [The development, validation, and reliability of short-forms of current instruments for the evaluation of therapeutic progress in psychotherapy and psychiatry]. *Diagnostica, 52*, 11–25. <http://dx.doi.org/10.1026/0012-1924.52.1.11>
- Lutz, W., Zimmermann, D., Müller, V. N. L. S., Deisenhofer, A. K., & Rubel, J. A. (2017). Randomized controlled trial to evaluate the effects of personalized prediction and adaptation tools on treatment outcome in outpatient psychotherapy: Study protocol. *BMC Psychiatry, 17*, 306–316. <http://dx.doi.org/10.1186/s12888-017-1464-2>
- 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: Theory, Research & Practice, 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>
- 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. <http://dx.doi.org/10.1037/0022-006X.68.3.438>
- Muran, J. C., Safran, J. D., & Eubanks-Carter, C. (2010). Developing therapist abilities to negotiate alliance ruptures. In J. Muran & J. Barber (Eds.), *The therapeutic alliance: An evidence-based guide to practice* (pp. 320–340). New York, NY: Guilford Press.
- Ogles, B. M. (2013). Measuring change in psychotherapy research. In M. J. Lambert (Ed.), *Bergin and Garfield’s handbook of psychotherapy and behavior change* (pp. 134–166). New York, NY: Wiley.
- Ormhaug, S. M., Shirk, S. R., & Wentzel-Larsen, T. (2015). Therapist and client perspectives on the alliance in the treatment of traumatized adolescents. *European Journal of Psychotraumatology, 6*, 27705. <http://dx.doi.org/10.3402/ejpt.v6.27705>
- Orth, U. (2013). How large are actor and partner effects of personality on relationship satisfaction? The importance of controlling for shared method variance. *Personality and Social Psychology Bulletin, 39*, 1359–1372. <http://dx.doi.org/10.1177/0146167213492429>
- Pepinsky, H. B., & Karst, T. (1964). Convergence: A phenomenon in counseling and in psychotherapy. *American Psychologist, 19*, 333–338. <http://dx.doi.org/10.1037/h0044396>
- Piper, W. E., & Ogrodniczuk, J. S. (2010). The therapeutic alliance in group therapy. In J. Muran & J. Barber (Eds.), *The therapeutic alliance: An evidence-based guide to practice* (pp. 263–282). New York, NY: Guilford Press.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Raue, P. J., & Goldfried, M. R. (1994). The therapeutic alliance in cognitive-behavior therapy. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: Theory, research and practice* (pp. 131–152). New York, NY: Wiley.
- Rozmarin, 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.
- Rubel, J. A., Rosenbaum, D., & Lutz, W. (2017). Patients’ in-session experiences and symptom change: Session-to-session effects on a within- and between-patient level. *Behaviour Research and Therapy, 90*, 58–66. <http://dx.doi.org/10.1016/j.brat.2016.12.007>
- Safran, J. D., & Muran, J. C. (1996). The resolution of ruptures in the therapeutic alliance. *Journal of Consulting and Clinical Psychology, 64*, 447–458. <http://dx.doi.org/10.1037/0022-006X.64.3.447>
- Safran, J. D., Muran, J. C., & Eubanks-Carter, C. (2011). Repairing alliance ruptures. *Psychotherapy: Theory, Research & Practice, 48*, 80–87. <http://dx.doi.org/10.1037/a0022140>
- Sasso, K. E., Strunk, D. R., Braun, J. D., DeRubeis, R. J., & Brotman, M. A. (2016). A re-examination of process-outcome relations in cognitive therapy for depression: Disaggregating within-patient and between-patient effects. *Psychotherapy Research, 26*, 387–398. <http://dx.doi.org/10.1080/10503307.2015.1026423>
- Shanock, L. R., Baran, B. E., Gentry, W. A., Pattison, S. C., & Heggstad, E. D. (2010). Polynomial regression with response surface analysis: A powerful approach for examining moderation and overcoming limitations of difference scores. *Journal of Business and Psychology, 25*, 543–554. <http://dx.doi.org/10.1007/s10869-010-9183-4>
- Sinclair, M., & Holmqvist, R. (2013). *Congruence in the therapists’ and clients’ ratings of the therapeutic alliance does not predict*

- outcome in psychological treatment. Retrieved from [www.coresverige.se/wp-content/uploads/2013/08/kongruens-i-alliansen-manuskript.docx](http://www.coresverige.se/wp-content/uploads/2013/08/kongruens-i-alliansen-manuskript.docx)
- Strauss, J. L., Hayes, A. M., Johnson, S. L., Newman, C. F., Brown, G. K., Barber, J. P., & Beck, A. T. (2006). Early alliance, alliance ruptures, and symptom change in a nonrandomized trial of cognitive therapy for avoidant and obsessive-compulsive personality disorders. *Journal of Consulting and Clinical Psychology, 74*, 337–345. <http://dx.doi.org/10.1037/0022-006X.74.2.337>
- Tasca, G. A., & Lampard, A. M. (2012). Reciprocal influence of alliance to the group and outcome in day treatment for eating disorders. *Journal of Counseling Psychology, 59*, 507–517. <http://dx.doi.org/10.1037/a0029947>
- 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>
- Tryon, G. S., Blackwell, S. C., & Hammel, E. F. (2007). A meta-analytic examination of client-therapist perspectives of the working alliance. *Psychotherapy Research, 17*, 629–642. <http://dx.doi.org/10.1080/10503300701320611>
- Wang, L. P., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods, 20*, 63–83. <http://dx.doi.org/10.1037/met0000030>
- West, T. V., & Kenny, D. A. (2011). The truth and bias model of judgment. *Psychological Review, 118*, 357–378. <http://dx.doi.org/10.1037/a0022936>
- Zandberg, L. J., Skriner, L. C., & Chu, B. C. (2015). Client-therapist alliance discrepancies and outcome in cognitive-behavioral therapy for youth anxiety. *Journal of Clinical Psychology, 71*, 313–322. <http://dx.doi.org/10.1002/jclp.22167>
- 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., Muran, J. C., Hungr, C., Eubanks, C. F., Safran, J. D., & Winston, A. (2016). The relationship between alliance and outcome: Analysis of a two-person perspective on alliance and session outcome. *Journal of Consulting and Clinical Psychology, 84*, 484–496. <http://dx.doi.org/10.1037/ccp0000058>
- Zilcha-Mano, S., Snyder, J., & Silberschatz, G. (2016). The effect of congruence in patient and therapist alliance on patient's symptomatic levels. *Psychotherapy Research, 27*, 371–380.
- 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>

Received September 20, 2016

Revision received September 5, 2017

Accepted November 24, 2017 ■

## Special Issue Call for Manuscripts on Therapist Transformations

**Editor:** Jennifer Callahan, PhD, ABPP

**Submission Deadline:** July 1, 2018

The *Journal of Psychotherapy Integration* announces a call for papers for a special issue on “**Therapist Transformations.**” The special issue will consist of pieces from practicing psychotherapists about the impact of and coping with illness or death of family members, life-threatening disease (e.g., cancer, heart attack, stroke) or disability, involvement in human-made and natural disasters, traumatic exposure, or significantly life altering experiences (e.g., adoption, divorce).

Manuscripts should draw out for others how personal challenges affected oneself as a psychotherapist and highlight what can be learned by others from these experiences, citing relevant works to ground the submission in scholarship.

We are seeking manuscripts of approximately 20 pages, not including references. All manuscripts should be blinded for peer review and submitted electronically by **July 1, 2018** through the journal's submission portal at [www.editorialmanager.com/pti/](http://www.editorialmanager.com/pti/).

Please read the Instructions to Authors ([www.apa.org/pubs/journals/int](http://www.apa.org/pubs/journals/int)) for more details on submitting manuscripts. Please submit manuscripts as a Special Issue article type, and note in your cover letter that you are submitting for consideration to be published in the special issue on therapist transformations.

For questions and consultations, please contact the Editor at [Jennifer.Callahan@unt.edu](mailto:Jennifer.Callahan@unt.edu).